

牌の有無や点棒の数等の原始的な特徴を用いた麻雀評価関数の性能解析

山本陽平 保木邦仁
電気通信大学

概要 麻雀において、牌の有無や点棒の数等の原始的な特徴を用いた三層ニューラルネットワークの性能を解析した。ニューラルネットワークは人間上級者の牌譜を教師として学習し、現局面におけるゲームの結果 (1 位から 4 位の着順) を予測した。順位予想の正答率は約 35% に達することが示され、予測は主に点棒の数によりなされていることが明らかとなった。また、危険牌やシャンテン数の認識も、ある程度の精度で成されることが示された。

Performance analysis of Mahjong evaluation functions represented by primitive features as tile positions and scores

Yohei Yamamoto and Kunihito Hoki
The University of Electro-Communications

Abstract: This work analyzed the performance of three-layer neural network with primitive inputs as tile positions. We trained the neural network by using game records of human experts to predict game results (1st, 2nd, 3rd, or 4th rank) from current game positions. It was shown that the percentage of correct prediction of the game results reaches about 35%, and the neural network predicts the game results mainly by means of current scores of four players. We also examined the accuracies of risk prediction to discard each tile and shanten number (the minimum number of tiles needed to reach tenpai) prediction.

を行い、性能評価をし結果を報告する。なお、本研究では評価関数に 3 層ニューラルネットワークを採用した。

1 はじめに

近年、思考対戦ゲームにおいてコンピュータが人間と同等または上回る成績を残したというニュースをよく耳にするようになった。チェスではコンピュータがハンデを負う対局において勝ち越し、第三回将棋電王戦ではコンピュータが 4 勝 1 敗という成績を残した。

不完全情報多人数ゲームにおける強いコンピュータの開発は、いまだ挑戦的な課題と考えられる。麻雀においては、これまでに、機械学習法を用いたゲーム状態の優劣評価に関する研究が数件報告されている。これらの先行研究では、面子や搭子の数や受け入れ枚数等のゲームの知識を評価関数の特徴に用いる。その一方で、牌の有無や点棒の数等のより原始的なゲームの構成要素のみを特徴とした評価関数の性能に関してはいまだ報告がない。本研究では、このような原始的な特徴からなる評価関数の機械学習

2 先行研究

北川らは評価関数として 3 層ニューラルネットワークを用いた教師あり学習で麻雀 AI のパラメータ調整を行った。入力に使用したゲーム特徴は、面前持ち牌 2 枚及び 3 枚の組み合わせや面子、搭子の数、安全度の高い牌等である。牌譜との一致率はツモ局面で約 56%、鳴き局面で約 89% となり、東風荘 [1] でのレートは 1318 となった [2]。水上らは牌譜との一致を目指した平均化パーセプトロンを用いて 1 人麻雀プレイヤーの特徴量の学習を行い、降りの技術を加して 4 人麻雀への適用をした。入力に使用したゲーム特徴は、シャンテン数や最も多い色の数、受け入れ枚数、タンヤオ牌の数等である。1 人麻雀プレイヤーの牌譜との一致率は 53% 程度となり、ネット麻

雀サーバ『天鳳』[3]でのレートは1507となった[4]。

3 実験について

麻雀の牌譜から得られる情報からニューラルネットワークの入力および出力を決定し学習を行い、その性能の考察を行う。本実験では複数のシグモイドユニットからなる3層ニューラルネットワークを用いた。3層ニューラルネットワークは入力層、中間層、出力層の3つの層で構成されており、入力層に入力を与えると、入力層から中間層、中間層から出力層の順で一方向に信号が伝播していく。学習は誤差逆伝播法を用いて、実際の出力値と理想の出力値の2乗誤差を最小化し、中間層と出力層の重みを調整、慣性項を付加した最急降下法により重みを更新することで行った。ネットワークの重みの更新は学習パターン1つを処理する度に行い、全ての学習パターンを処理して一回の反復とした。重みの初期化や学習パターンの並びに関わる乱数のシード値は1から12までとしそれぞれデータを取得した。また、学習定数は0.01、安定化定数は0.8とした。なお、結果に用いた図のプロット点はシード値を変えた結果の平均、エラーバーは標準誤差の2倍としている。横軸が反復回数の図は横軸を対数でとり、1の時を学習なしの状態としている。

以下では初めに、対局終了時順位を予測する実験結果を示す。続いて、打牌後の放銃予測、またシャンテン数認識の実験結果を示す。

3.1 実験 1-1 出力:対局終了時順位

麻雀の牌譜から学習パターンとテストパターンを生成した。パターンは入力信号と出力信号の組からなる。入力層のユニット数は1343個で、入力信号の内訳及び符号化の例は表1の通りである。なお、牌を表現する場合は常にマンズ・ピンズ・ソーズ・字牌の順で並んでおり、下家、対面、上家は直前の選択プレイヤーから見た位置関係を表している。出力層のユニット数は1個、中間層のユニット数は672個とした。出力信号は対局終了時順位に対応し、1位を1/8、2位を3/8、3位を5/8、4位を7/8とした。順位予測は、出力値が0から1/4の値で1位、1/4から1/2の値で2位、1/2から3/4の値で3位、3/4から1の値で4位とした。なお符号化した局面はシステムティック麻雀研究所[5]にて公開されている、東風荘[1]における「とつげき東北」氏及び「菌太郎」氏の牌譜のうち東風戦のもので打牌直後の局面を用いている。学習の反復回数は20回で打ち切りとし、学習パターン数は1万、2万、4万、8万、16万、32万、64万、テストパターン数は1万とした。

3.1.1 結果 1-1

図1は反復回数に対する平均2乗誤差の推移のうち、学習信号数が1万、4万、16万、64万のものを

示した図である。横軸が反復回数、縦軸が平均2乗誤差であり、右上の数値は学習パターン数と各線の対応である。この結果より、学習が進むと共に平均2乗誤差が小さくなっていることは分かるが、反復回数20回の段階で収束には至っていないと考えられる。

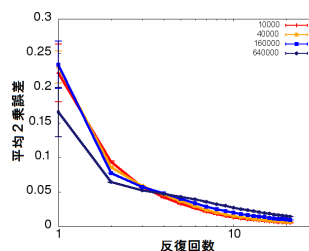


図1 実験 1-1: 反復回数と平均2乗誤差

図2は反復回数に対する対局終了時順位予測の完全一致率である。横軸が反復回数、縦軸が順位予測の完全一致率であり、右下の数値は学習パターン数と各線の対応である。テストパターンでの予測順位と実際の順位との一致率は約35%に達し、無作意に予測した場合の一致率(25%)より高い値が得られた。重みの初期値や学習パターンの並びをかえると、約1%程度一致率が変化した。また、学習パターン数が2万以下では一致率の低下が確認された。

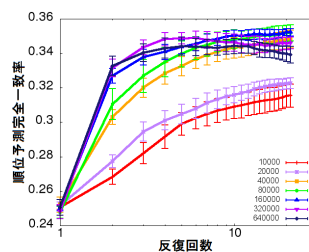


図2 実験 1-1: 反復回数と順位予測完全一致率

図3は実際の打牌が出力値のランクの何位までに含まれているかを示した図である。出力値のランクは、各局面で全通りの打牌を行い、打牌後の局面を入力して出力値を計算し得た。使用した局面はリーチ後の打牌選択を行うことが出来ない局面以外の打牌局面を用いた。横軸が出力値順位 n 、縦軸が n 位までの打牌一致率となっている。学習を行っていないニューラルネットワークを用いてランダムな打牌の結果も含めた。この結果より、打牌一致率は約10%であり、出力値を対局終了時順位とした場合、学習後でもランダムとほぼ変わらない打牌選択を行うということが分かった。また、北川ら[2]の56%や

表1 入力層のユニット内訳及び符号化の例

内容	個数	補足説明	例
局数	4個	東1局、東2局、東3局、東4局以降	東2局 - 0100
本場	4個	0本場、1本場、2本場、3本場以降	2本場 - 0010
積みリーチ棒	3個	0本、1本、2本以上	2本 - 001
表ドラ	34個	同じドラとなる場合は1個とカウント	五萬 - 00001000...000
直前選択プレイヤーの風	4個	東、南、西、北	東 - 1000
各プレイヤーの点数	51*4個	持ち点1000で0から50(百以下切り捨て、50以上は50)	該当する点数の入力ユニットに1
直前選択プレイヤーの面前牌	136個	副露牌は含まない	一萬*3、二萬*1、三萬*2のみ所持 - 1110100011000000...0000
直前選択プレイヤーの副露牌	(4+34)個	(副露種:4 + 副露牌の先頭の牌:34)*4 副露はボン、チー、暗カン、明カンの順	五萬を暗カン - 001000001000...000
直前選択プレイヤーに通る牌	34個	直前選択プレイヤーの切った牌と 直前選択プレイヤーがリーチ後に他のプレイヤーに切られた牌	割愛
直前選択プレイヤーのリーチの有無	1個	リーチをしていれば1	割愛
下家、対面、上家の副露牌、通る牌、リーチの有無	187*3個	下家(副露牌、通る牌、リーチの有無)、対面...の順 表現は直前選択プレイヤーの時と同様	割愛
直前選択プレイヤーにとって不可視の牌	136個	不可視の牌を1	一萬*3、二萬*1、三萬*2のみ可視 - 1000111011000000...0000
ツモ可能な残りの山の枚数	70個	0枚から69枚	残り2枚 - 001000...000

水上ら [4] の 53% には及ばない結果となった。

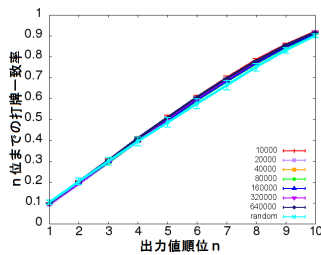


図3 実験 1-1: 実際の打牌との出力値順位一致率

3.2.1 結果 1-2

図4は反復回数に対する対局終了時順位予測の完全一致率である。学習パターン数に関わらず35%から36%の間に収束しているように見える。また、表2に直前の選択プレイヤーの持ち点と出力値との相関係数をまとめた。この結果より、学習パターン数が多くなるほど強い相関があることが分かった。

表2 直前の選択プレイヤーの持ち点と出力値の相関係数

学習パターン数	相関係数
1万	-0.740 ±0.006
2万	-0.808 ±0.008
4万	-0.869 ±0.002
8万	-0.902 ±0.004
16万	-0.915 ±0.003
32万	-0.907 ±0.002
64万	-0.910 ±0.001

3.2 実験 1-2 入力:各プレイヤーの点棒 出力:対局終了時順位

実験 1-1 の結果を受け、出力などの条件は変えず、入力に「各プレイヤーの点棒」のみ与え学習を行った。順位予測の一致率と直前選択プレイヤーの持ち点と出力値との相関から実験 1-1 の結果について考察する。

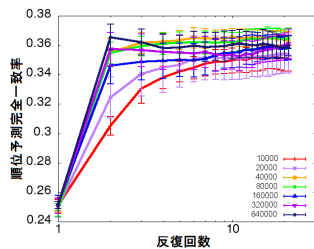


図4 実験 1-2: 反復回数と順位予測完全一致率

順位予測の完全一致率が実験 1-1 と 1% 程度しか変わらないこと、また持ち点と出力値に強い相関があることを考えると、出力を対局終了時順位とした場合、持ち点が多い程順位が良いと予測をしているに過ぎないと考えられる。

3.3 実験 2 出力:入力局面の放銃の有無

入力層や使用した牌譜、学習及びテストパターン数等を実験 1-1 と同様、出力を入力局面における放銃の有無として牌の危険度を学習することが出来るかを、実験 1-1 と同様に乱数のシード値を変化させて評価を行った。

3.3.1 結果 2

放銃予測の完全一致率は全ての学習パターン数において約 99% となった。これは放銃した局面に対して、放銃していない局面があまりに多いため、全てにおいて放銃しないと予測したため、一致しなかった分だけ放銃局面があったことを意味している。

64万パターンで学習したニューラルネットワークに入力を与え、各シード毎に実際の出力値の各区間ごとの放銃の頻度を計算し、その平均と信頼区間を表3にまとめた。ニューラルネットワークには出力値0.25を放銃なしとして学習させている。この表3より、0.27から0.32までの区間において、出力値0.23から0.26までの区間よりも放銃率に有意差があることが分かる。この結果より、3層ニューラルネットワークで危険牌の予測がある程度はできていると考えられる。

表3 出力値の各区間における放銃頻度

出力値	放銃率 %	
0.23~0.24	1.8	±0.7
0.24~0.25	0.43	±0.05
0.25~0.26	0.42	±0.02
0.26~0.27	1.29	±0.04
0.27~0.28	2.7	±0.1
0.28~0.29	3.7	±0.2
0.29~0.30	4.9	±0.5
0.30~0.31	5	±1
0.31~0.32	6	±3
0.32~0.33	5	±5

3.4 実験3 出力:シャンテン数

入力を「直前選択プレイヤーの面前牌」と「直前選択プレイヤーの副露牌」として3層ニューラルネットワークのシャンテン数認識の精度を検証した。出力層および中間層の数は実験1-1と同様、学習パターン数を1万、2万、4万、8万、16万、テストパターン数を1万として反復回数を20回で打ち切りとして学習を行った。実験1-1と同様に乱数のシード値を変化させて評価を行った。出力信号では、0から8のシャンテン数を1/18, 3/18, ..., 17/18の数値で表現した。シャンテン数予測は、出力値が0から1/9の値で0、1/9から2/9の値で1などとした。但し、ここで扱うシャンテン数は、4面子1雀頭のアガリ形に対するシャンテン数とした。

3.4.1 結果3

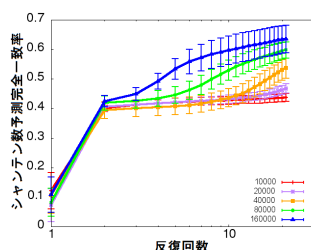


図5 実験3: 反復回数とシャンテン数認識完全一致率

図5は反復回数に対するシャンテン数予測の完全一致率である。横軸が反復回数、縦軸がシャンテン数予測の完全一致率であり、右下の数値は学習パターン数と各線の対応である。学習パターン数が16万、

反復回数20回の時、一致率は約64%となった。また、図5の推移から、反復回数を増やすことでさらに一致率は上昇するものと考えられる。しかしながら、実際のプレイにおいてシャンテン数は手牌を知ることが出来れば正確に知ることが出来る。シャンテン数が正しく認識できないと、ノーテンでリーチしたり、テンパイでリーチできなかったりするため、本実験で得られたニューラルネットワークはゲームをプレイするのに十分な認識精度をもたない。

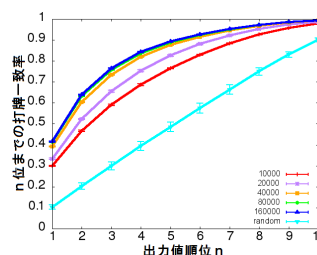


図6 実験3: 実際の打牌との出力値順位一致率

図6は実際の打牌が出力値のランクの何位までに含まれているかを示した図である。16万パターンで学習を行った際の打牌一致率は約42%と、出力値を対局終了時順位としたときよりも一致率は約30%上昇した。しかし、先行研究の一致率と比較するとまだ10%もの開きがあり、今回用いた特徴を持つ評価関数は良い性能とは言えないだろう。

4 おわりに

本研究では、牌の有無や点棒の数等のより原始的なゲームの構成要素のみを特徴とした評価関数の性能について考察を行った。それにより、3層ニューラルネットワークが持ち点が多い程最終順位が良くなるというゲームの性質を認識している事が確認された。また、牌の危険度やシャンテン数についても、ある程度認識している事も確認された。しかしながら、シャンテン数の予測精度は約64%程度であり、牌の位置情報等の原始的なゲーム特徴のみを入力した3層ニューラルネットワークをもちいてAIを構築することは困難であると考えられる。

参考文献

- [1] 東風荘 <http://mj.giganet.net/>
- [2] 北川竜平, 三輪誠, 近山隆, 麻雀の牌譜からの打ち手評価関数の学習, Game Programming Workshop 2007, pp.76 - pp.83, Nov.2007
- [3] 天鳳 <http://tenhou.net/>
- [4] 水上直紀, 中張遼太郎, 浦晃, 三輪誠, 鶴岡慶雅, 近山隆, 降りるべき局面の認識による1人麻雀プレイヤーの4人麻雀への適用, Game Programming Workshop 2013, Nov.2013
- [5] とつげき東北, システムティック麻雀研究所 <http://totutohoku.b23.coreserver.jp/hp/>

第9回 E&C シンポジウムに投稿された本論文の著作権は、著者にあります。
著者に無断の複製は厳禁です。複製などのご希望は、著者に直接御連絡下さい。

エンターテイメントと認知科学研究ステーション

代表連絡先

〒182-8585 調布市調布ヶ丘1-5-1

電気通信大学 情報理工学研究科 伊藤毅志

uec-ito@mbc.nifty.com

042-443-5370