

ニューラルネットワークを用いた 麻雀の捨て牌危険度における過学習の影響

矢ノ口裕貴[†] 篠埜功[‡]

芝浦工業大学^{†‡}

1. はじめに

麻雀では、一般的に1ゲームにつき複数の試合を行い、その終了時点での点数の大小で順位を決定するため、最終的な得点の増大のために、他のプレイヤーのあがりとなる牌を捨てないことにより失点を抑えることは重要な戦略となる。手配の特定の牌が他のプレイヤーのあがりとなる牌であるかの指標を牌の危険度とすると、危険度の高い牌を推定することが出来れば、その牌を捨てないことにより失点を抑えることが出来る。関連研究として、牌を捨てたときに失う点数の期待値の推定をSVRによって行った我妻ら[1]の研究がある。本研究では、プレイヤーが捨てようとしている牌が他のプレイヤーのあがりとなる牌であるかを推定することを目的として、ニューラルネットワークの出力値を牌の危険度として用いることとする。ニューラルネットワークによって学習を行うに際して、過学習の影響がどの程度見られるのかについても調査した。

2. 牌譜学習

本研究では、実際の牌譜からあるプレイヤーが他のプレイヤーにあがりとなる牌を捨てた局面を構築し、そこから評価要素を抽出してニューラルネットワークに入力することで、その出力値が牌の危険度になるような学習を行う。

2.1 学習データ

学習データとして、とつげき東北システムティック麻雀研究所[2]にて公開されている牌譜を用いる。牌譜とは、各プレイヤーの初期手牌、引いてきた牌、捨てた牌などの記録であり、それらの情報から各局面を構築して評価要素を抽出し、ニューラルネットワークへの入力とする。

2.2 誤差修正

牌譜内の各局面における合法手による次の局面の評価値と、牌譜内での実際の手による次の局面の評価値との差を求め、誤差逆伝播法によりパラメータ

の調整をすることでニューラルネットワークのノード間の重みを更新する。この方法は北川ら[3]の手法を参考にしているが、本研究では局面から抽出する評価要素を変更し、また、学習する局面を2.4節で述べるように限定している。

ニューラルネットワークの重みベクトル ω の更新は以下のように求められる。

$$\omega_{new} = \omega_{old} - \eta \Delta \omega_{old}$$

ω_{old} は更新前の重みベクトル、 ω_{new} は更新後の重みベクトル、 η は学習率である。

重みベクトルの更新は、全合法手についての牌譜の手との評価値の差をシグモイド関数によって求め、それらの和を最小化させるように行われる。このような学習を繰り返すことによって得られたニューラルネットワークは、学習に用いた局面と似た局面の評価値が高くなり、出力値として牌の危険度を出力することが期待される。

2.3 過学習

学習方法に誤差逆伝播法を用いる際には、過学習現象が問題となる場合がある。過学習とは、学習データに適合するように学習されたニューラルネットワークを用いてテストデータを評価するときに、意図した評価基準を満たさない現象のことをいう[4]。本研究では、ニューラルネットワークの中間層の数や学習率を変動させることで、汎化性能を向上させることが出来るか調べた。

2.4 学習方法

はじめに、牌譜から学習する局面を構築する。本研究では学習する局面を、いずれかのプレイヤーがあがりに成功した試合のうち、ほかのプレイヤーがあがりとなる牌を捨てた局面に絞り、あがりとなる牌を捨てたプレイヤーの視点から見る事が出来る牌のみを評価要素として使用した。

本研究では、以下の手順で学習を行った。

- (1) 学習局面として抽出した各局面から、牌譜上の手によって実際に移行した局面と、合法手によって移行することが出来る局面をそれぞれ構築する。

Overfitting in Risk Estimation Associated with Discarding Tiles in Mahjong by Neural Network

[†]Yanokuchi Yuki, Shibaura Institute of Technology

[‡]Sasano Isao, Shibaura Institute of Technology

(2) 構築した局面から評価要素を抽出する。

(3) 抽出した評価要素をニューラルネットワークに入力して評価値を計算し、牌譜で実際に移行した局面と合法手で移行できる局面との評価の誤差を計算する。得られた誤差をもとに勾配の計算を行い重みベクトルの更新をする。

(1)~(3)の手順を学習する全局面に対して行う。これをあらかじめ設定した学習回数分だけ繰り返し行う。

2. 5 学習結果の評価

牌譜内の各局面において、各合法手によって移行できる局面に対するニューラルネットワークの評価値のうち、最上位のものが牌譜内の手と一致する割合を完全一致率、上位3つのうちいずれかと一致する割合を3位以内一致率とし、学習の精度を図るための尺度とする。

牌譜による学習を終えた後、ニューラルネットワークに対して学習に用いていない牌譜をテスト局面として入力し、テスト局面とニューラルネットワークの出力値の一致率を求めることにより牌の危険度推定の性能を評価する。麻雀の試合においては、牌を捨てても他のプレイヤーのあがりにならない局面が多いため、牌の危険度を自動で評価するために、テスト局面として牌譜からあがりとなる牌を捨てた局面を抽出して使用する。

3. 実験

Java 言語でプログラムを実装し、試合数 5000、1局面当たりの学習回数を 10 回とし、ニューラルネットワークの中間層の数を 50, 200、学習率のパラメータを 0.9, 0.8, 0.7, 0.6, 0.5 と変化させ学習した。ニューラルネットワークへの入力とする評価要素および実験結果を以下に示す。

3. 1 評価要素の設定

牌譜から学習する各局面を構築して評価要素を抽出し、ニューラルネットワークへの入力とする。局面の評価要素は我妻ら[1]の研究を参考に選出しているが、牌同士のつながりに関する評価要素を中心として設定し、あがり役に関する評価要素は必要最小限のものに留めている。また、あがったプレイヤーと自プレイヤーの牌の動きに注目するため、他のプレイヤーの牌の動きに関する評価要素は設定していない。

3. 2 実験結果

5000 試合分の牌譜内から学習する局面として抽出した局面は 3215 局面である。牌譜学習を行った後、ニューラルネットワークに対してテスト局面を入力し、それぞれの局面に対して評価値の高い着手上位3つと牌譜内の手との一致率を調べた。テスト

局面として抽出した局面は 750 局面である。実験結果の一部を表 1 に示す。

表 1: 中間層 50 個、200 個での学習率ごとのテスト局面との 3 位以内一致率

	中間層 50 個	中間層 200 個
学習率 0.9	22.53%	22.00%
学習率 0.8	21.60%	22.13%
学習率 0.7	22.40%	22.00%
学習率 0.6	22.53%	22.27%
学習率 0.5	22.53%	22.13%

実験では、ニューラルネットワークの中間層の数や学習率の変化によって、テスト局面との一致率にほとんど差異が見られなかった。

4. まとめと今後の課題

本研究では、ニューラルネットワークの機械学習により捨て牌の危険度を推定することを目的とし、実際の人間のプレイヤーによる牌譜をもとに学習を行った。この際に、ニューラルネットワークの中間層の数や学習率を変化させることで、ニューラルネットワークの汎化性能が向上するか調査した。学習によって得られた危険度推定の評価として、テスト局面との一致率を比較した。実験により、中間層の数や学習率に関わらず、テスト局面に対して平均で約 22.2%の一致率を得ることが出来た。

今後の課題としては、学習する局面数や局面の評価要素を増やすことにより、推定の精度を上げることが挙げられる。また、テスト局面との一致率が学習回数や中間層の数によってほとんど変化しなかったため、その理由について探っていくことが挙げられる。牌の危険度の評価方法としては、本研究ではあがり局面においてプレイヤーが捨てたあがり牌と一致するかどうかによって評価したため、あがり局面以外でも牌の危険度の妥当性を客観的に評価できる方法を考案することが挙げられる。

参考文献

- [1] 我妻敦, 原田将旗, 森田一, 古宮嘉那子, 小谷善行: SVR を用いた麻雀における酔手牌の危険度の推定, 情報処理学会研究報告ゲーム情報学, 2014-GI-31(12), pp. 1-3, 2014.
- [2] 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/>.
- [3] 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, ゲームプログラミングワークショップ 2007 論文集, pp. 76-83, 2007.
- [4] 小川英光, 山崎一孝: 過学習の理論, 電子情報通信学会論文誌 D, vol. J76-D-II, no. 6, pp. 1280-1288, 1993.