

人間プレイヤーの戦略を推定し、 系列情報を用いて人間の模倣を行う麻雀プレイヤーの作成

前川 幸輝¹ 竹内 聖悟²

概要：4人で行う不完全情報ゲームである麻雀は、次の牌を予想することができないため、人間プレイヤーは様々な未来の手牌を想像して現在の打牌を選択する。また、手牌の構成や持ち点の状況を含めると、全く同じ局面はほとんど現れないため、打牌選択に際して人間プレイヤーが取る戦略も多様である。

ところで、近年は「オンライン麻雀」という形態でインターネットを通じて、世界中の人々と対戦が可能である。しかし、オンライン麻雀では環境を理由として回線が切断されたプレイヤーが意思のない打牌を続け、対局に影響を及ぼす点が課題である。本研究は、この課題を解決するために、人間プレイヤーの打牌を記録した「系列情報」を用いて、人間プレイヤーに代わって打牌を模倣する麻雀プレイヤーを提案する。

Estimating Strategy and Imitating Play Style of Human Players Using Sequence Information in Mahjong

KOKI MAEKAWA¹ SHOGO TAKEUCHI²

Abstract: Mahjong, a four-player incomplete information game, does not allow the player to predict the next tile, so the human players try to imagine various future hands and select the current tile to play.

In addition, there are many strategies that a human player can use when selecting his or her tiles, because in mahjong, there are few situations in which the tiles are exactly the same, including the composition of the hand and the difference in scores.

By the way, in recent years, it has become possible to play "online mahjong" via the Internet, where you can play against people from all over the world.

However, the problem with online mahjong is that the line can be disconnected, and the player will continue to play tiles that they haven't considered, which will give a negative impact on the game.

In this research, we propose a mahjong player that imitates a human player by using "sequence information" that records the tiles played by the players and tries to solve this problem.

1. はじめに

麻雀は4人で行うボードゲームであり、不完全情報多人数ゲームに分類される。対局者は「山」という無作為に積まれた牌の集合から、毎回の自分の手番に抽選を行い牌を得るが、山の牌は伏せられているため、次に得る牌を事前には予想できない。このため、対局者は目標とする牌姿を想像しながら、自分の手番で欲しい牌を1枚手に加え、代わりに不要となった牌を場に1枚捨てることを繰り返して和了を目指す。

麻雀は和了によって点棒をやり取りし、最終局の終了時に点数を持っている対局者から1位、2位、3位、4位と順位が定まる。しかし、1位になるという目標を全員が共有した対局においても、手の進め方や行動のような「戦略」は、人間プレイヤーのそれぞれが持つ基準により、重視する点が異なる。また、使用する牌は34種136枚存在し、「鳴き」と呼ばれる特殊な行動や、対局者の持つ点数の状況によっても変化するため、全く同じ局面が二度以上現れることは事実上ない。

麻雀は世界各地に拡がり楽しまれているため、多様なルールが存在し、標準化された絶対的なルールというもの

¹ 高知工科大学大学院 工学研究科

² 高知工科大学大学 情報学群

はない。ところが、近年はインターネットを通じて、共通のルールで世界中の人々と対戦可能な「オンライン麻雀」が提供されている。中でもオンライン麻雀の1つである「天鳳^{*1}」は、対局の詳細な記録である「牌譜」が公開され、再現性に優れているため、様々な学術的な研究で活用されてきた。

ただし、オンライン麻雀での対局は、依存するインターネット環境により、回線が切断される点が課題である。その場合、回線切れを起こした対局者は、意志がない打牌を繰り返す状態となり、対局に影響を及ぼす。そこで、本研究では人間プレイヤーの戦略と打牌を重視しつつ、人間プレイヤーに代わってその打牌選択を模倣する麻雀プレイヤーを提案する。

2. 関連研究

麻雀の対局者の牌譜を学習に用いて、強い麻雀プレイヤーを作成しようとする試みは数多く行われてきた。水上らは上級者の牌譜をパーセプトロン学習して、1人麻雀プレイヤーを作成し、それを拡張することで、平均的な人間プレイヤーよりも強い実力のコンピュータプレイヤーを作成した [1]。水上らは「鳴き」の学習のために、攻めていると考えられる鳴きが発生した局面だけを選ぶことで、元の牌譜との正解率が 84.2% に達した。

一方で、人間の牌譜を用いず、シミュレーションによって対局者の最適な行動を探る研究も存在する。吉村らはタブーサーチアルゴリズムを採用して、初期手牌から分岐し得る最適な和了形を探索した。打牌局面と鳴き局面の両方で探索を行い、牌譜と提案手法との打牌の一致率により評価する実験の結果、最大で 85% の打牌が一致する麻雀プレイヤーを得た [2]。

堺田らは麻雀上級者の捨て牌に対して、頻出するパターンをシーケンシャルパターンマイニングによって抽出し、ロン和了をした勝者と、振り込んだ敗者の捨て牌のパターンの違いを明らかにした [3]。

Kurita らは、ディープニューラルネットワークを用いずに、マルコフ決定過程 (MDP) を導入した木探索によるシナリオを構築し、向聴数によってそのシナリオを切り替える麻雀プレイヤーを作成し、人間の上位 200 位に入るレベルに達したと結論付けた [4]。

本研究では、対局者の手牌に着目して打牌のパターンから系列を取り出し、その打牌を模倣することを目指す。

3. 麻雀の用語とルール

麻雀は4人で行うボードゲームであり、全ての対局者に公開されている部分と、それぞれの1人の対局者のみが視認できる非公開の部分が存在する不完全情報ゲームである。

麻雀における対局者の目的は、初めに均等に配られる点数を増やし、最終的に他のプレイヤーよりも多くの点数を保持しておくことである。点数を増やすには、自分の「和了」の成立による点数の受領、または、和了りが発生せずに終局した「流局」という状況で聴牌し、他のプレイヤーから罰符という形で点数を受け取る2種類の方法がある。

麻雀の進行は、開始時に与えられた13枚の手牌に1枚ずつモって1枚切ることを繰り返し、定められた「役」の形が成立した和了の状態となるまで手牌の1枚を交換する。和了は役の有る14枚の手牌を誰よりも早く揃えた者だけが宣言できる。

3.1 本稿で用いる麻雀に関する用語

本論文で使用する中で、解説が必要な用語を紹介する。

山 麻雀牌が無作為に二牌ずつ積まれたもの。

河 各対局者が捨てた牌を順番に並べていく場所。

自摸 山から牌を1枚取って手牌に加えること。

打牌 手牌から1枚の牌を河へと切り出すこと。

面子 手牌を分割する単位で、同じ種類の牌を3枚揃える刻子、4枚揃える槓子、同じ種類の連続する数字の3牌を揃える順子がある。

和了 対局者が手役の条件を満たし、それを宣言した状態のことである。和了には難易度に応じた点数が設定されており、和了の宣言を受けた対局者は規定の点数を支払わなければならない。

聴牌 あと1牌で和了が成立する状態のこと。

向聴数 最速の聴牌までに必要な牌の数を示す指標で、聴牌を0向聴と考え、数値が大きいくほど聴牌から遠い状態を表す。

立直 対局者が聴牌時に宣言できる行為で、発声してから1000点を場に供託する。立直を宣言して和了すると、「立直」という役が付き、「一発」や「裏ドラ」といった更なる打点上昇のチャンスがある代わりに、立直後は手牌構成を変えることができない。

即立直 立直の内、聴牌したその瞬間に宣言する立直のこと。

鳴き 他の対局者が捨てた牌を自分の手牌に組み入れる行為のこと。槓の場合に限り、相手の牌を使用しない鳴きもある。鳴いた場合には成立する役が減り、暗槓を除く鳴きをしたケースで立直ができなくなる。

3.2 オンライン麻雀「天鳳」について

「天鳳」は、2006年8月よりサービスを開始した大手オンライン麻雀ゲームで、累計ユーザー数が550万人を超える。対局者の強さを示す指標として、級位段位とレーティングが採用されており、対局者のレベルによって打つことができる卓に制限がある。段位戦は、初期登録時から打てる「一般卓」から順に、「上級卓」「特上卓」「鳳凰卓」が存在し、鳳凰卓からは段位とレーティングによる制限を満た

*1 <https://tenhou.net/>(2022年07月18日閲覧)

すことに加えて、月毎に定められた入場料を支払う必要がある。特上卓までは牌譜の URL の公開を条件に AI の参加も認められているが、鳳凰卓で麻雀を打つことは禁止されている。

3.2.1 「天鳳」における級位、段位について

級位は「新人」に始まり「1級」までの10段階、段位は「初段」に始まり「天鳳位」までの11段階が存在する。「1級」までは降級しないが「初段」からは、対局の順位に対応する「ポイント(Pt)」の配分によって降級または降段が存在する。上の段位に上がるためのポイントの要求量は、上位であるほど増えていく。なお、最高位に相当する天鳳位では降段は無い。

4. 提案手法

本研究では、ある対局者の打牌を模倣するために図1のモデルを提案し、構築する。

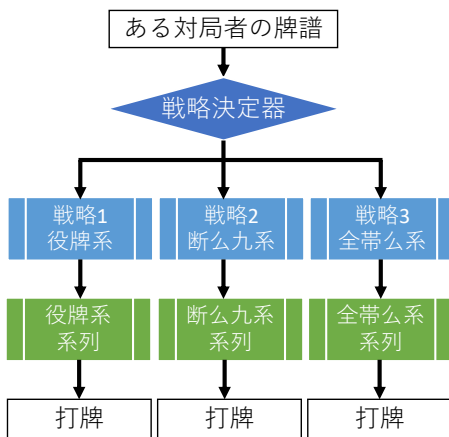


図1 対局者を模倣する麻雀プレイヤーのモデル

4.1 系列について

このモデルでは「系列」という概念を新しく定義するため、初めに系列について説明する。

前状態



次状態



図2 系列の具体例

系列は、タプル(前系列,次系列)から構成される。前系列は、自摸局面における対局者の手牌の数字の並びを色別に取り出したもので、次系列は打牌局面で取り出したものである。前系列から1枚を打牌して次系列となるため、次系列は対応する前系列よりも長さが常に1少ない。

図2は図中の「前状態」は索子の5の自摸局面、「次状態」は筒子の6の打牌局面であり、ここで得られる系列は

(3336678,333678)のみである。前状態では、筒子の3336678と、索子の33345を前系列として抽出する。系列を考える際には色情報を消し、数字のみを抜き出すため、字牌の「南南」は系列としては扱わない。なぜなら、実際の対局者は、色情報よりも牌の数字の組み合わせや、残り枚数を重視して打牌を選択していると考えられるためである。次状態でも同様に考え、次系列として筒子から333678、索子から33345を取り出すことができる。ここで、前系列と次系列が同じ形の場合は、系列として扱わないとする。つまり(33345, 33345)は系列に含まない。すなわち、図2から得られる系列は(3336678,333678)である。ただし、対局者が鳴いた場合には、他のプレイヤーから鳴きで得た牌を自摸牌とみなし、同様に系列を定義している。例えば、34が手牌にある際に2をチーした場合には、2を自摸牌とみなして、(34, 234)という系列を抽出する。

系列は、対局者の個々の局面における手牌の変化を表現している。このため、ある手牌を与えられた対局者が取る行動を、「人間に理解できる形」で個別に定義できるのが特徴である。手牌の和了への近さを表す一般的な指標には、向聴数が存在する。麻雀での手牌変化には123という形で4を引いて1を切り、234にする「スライド」と呼ばれる操作がある、この「スライド」の際には、向聴数は変化しないが、系列であればこの局面でも抽出が可能であるという利点もある。2019年の牌譜において、和了が成立した局面に限定した場合、(1234,123)は925回、(1234,234)は4626回観測された。

4.2 対局者を模倣するプレイヤーの作成

本研究の最終的な目標は、任意の局面で対局者を模倣し同じ打牌を出力する模倣プレイヤーを作成することであり、図1のモデルで表した。

- (1) 任意の毎局面で、牌譜の形式である局面の手牌を戦略決定器に与える。
- (2) ヒューリスティックに作成した戦略決定器を手牌に適用して、3種の中からもっとも到達可能性の高い手役へ向かう戦略を決定する。
- (3) 手牌の萬子、筒子、索子の数字部分だけを取り出し、前系列を求める。
- (4) 手牌の前系列から変化し得る全ての次系列に対して、過去の対局者のデータを探索し、その出現回数を比較することで、頻出する系列を求める。
- (5) 頻出する系列の一覧から出現回数の多い順に不要牌を複数候補に挙げ、過去に最も選択された系列と同じ形になるように打牌を優先的に選ぶ。この時、どの色を選ぶかという点については、その色の系列の出現数が最も多いものを採用する。

個々の対局者の系列は、対局者の天鳳における牌譜を用いて取得と解析をする。最終的な評価には、模倣プレイヤー

の構築には用いていない検証用データの牌譜を入力したときの出力と、牌譜の対局者の打牌との一致率を用いる。

5. 事前実験

系列を用いることが効果的な局面を探るために事前実験を行った。天鳳の牌譜で、和了が成立した局面において、成立した役と系列の関係を調べた。

5.1 実験で用いる牌譜データ

この節では、以後の実験で用いる牌譜データを説明する。役毎に系列を作成するため、牌譜を役別にまとめる作業を行った結果が表1である。表1において、「年」とは対局が開始された年である。「断タン九キウ成立」の列では、半荘を構成する1局の和了役の中に、断タン九キウが含まれていた局の数を示している。「全帯テン系成立」の列では、同様の1局の和了役の中に、混全帯テン九キウ（以降、単に全帯テンと表す）または純全帯テン九キウが含まれていた局の数を示している。全帯テン系キョウの役には他にも混老頭コンラウドや清老頭セイラウドがあるが、これらは出現数が極端に少ないため、データには加えていない。「役牌成立」の列も同様に役牌が含まれていた局の数を表しているが、役牌は全帯テンと複合する可能性があり、ここでは2回の重複を許して同じ局を計上している。

なお、以降ではこのデータの内、2012年から2018年を「系列の学習データ」と呼び、2019年を「系列の検証データ」と呼ぶ。

表1 特定の役が成立した局面数の比較

年	断 <small>タン</small> 九 <small>キウ</small> 成立	全帯 <small>テン</small> 系成立	役牌成立	合計
2012	69,449	3,410	102,602	175,461
2013	86,477	4,159	127,758	218,394
2014	90,017	4,525	135,309	229,851
2015	95,011	4,680	142,824	242,515
2016	89,476	4,554	134,797	228,827
2017	100,639	5,098	149,014	254,751
2018	100,616	5,094	149,265	254,975
2019	89,559	4,397	131,957	225,913

5.2 役による系列の差を測る実験と結果

2018年の鳳凰卓における特定の役が成立した対局の牌譜を用いて、役による出現系列の差を調査する実験を行った。ここでは手牌の全ての牌が2から8の数牌である「断タン九キウ」と、逆の全ての面子構成が1または9に絡み、字牌が必要となる「全帯テン」に着目した。全ての対局者に対し、断タン九キウと全帯テンが成立した局面で最も多く現れた系列を比較したところ、表2に示す結果となった。

表2 前系列が2枚以上の成立役別上位最頻系列

断 <small>タン</small> 九 <small>キウ</small> 成立系列	出現数	全帯 <small>テン</small> 成立系列	出現数
(6789, 678)	2494	(2789, 789)	354
(4678, 678)	2492	(6789, 789)	353
(1234, 234)	2461	(1789, 789)	346
(2346, 234)	2371	(1239, 123)	343
(346, 34)	2268	(1237, 123)	341
(467, 67)	2192	(1238, 123)	334

この結果、和了へ向かう系列には規則性があり、役に強く依存することが読み取れる。また、2から8の牌を自由に構成できる断タン九キウに対して、全帯テンでは構成可能な面子が123, 789, 111, 999のいずれかに限られるため、頻出する系列の関連牌がより強く偏った。以上より、対局者が特定の手役へ向かっている際の手順には頻出する傾向があり、その差は系列によって表すことができることが示された。

5.3 過去に出現した手牌との完全一致による手法との比較

本研究の提案手法と同様に、過去のデータから人間プレイヤーの次の一手を予測する手法には、データベースを用いた方法がある。つまり、過去に出現した手牌を全てデータベースに記憶しておき、毎局面でパターンマッチを行右ことで模倣を実現しようとするものである。

しかし、結論から言えばこれは失敗した。麻雀は手牌だけでも組み合わせが非常に多く、同じ局面はほとんど現れなかった。プログラムにより、2018年の全ての牌譜の手牌を辞書に登録してデータベースを構築し、2019年の全ての牌譜（およそ3320万局面）の手牌と比較した結果、完全一致したのは1件であった。同様に、2013年の手牌（およそ3120万局面）と比較した場合でも完全一致したのは1件であった。ただし、この1件については、非常に大きな数を取り扱う中で発現した誤差の可能性もある。

なお、系列では、手牌の数字列に着目し、更に色情報を落としていることで、完全一致がほとんどないという問題を緩和している。

6. 実験

実験は、提案手法では、個々の対局者の打牌選択を模倣する代わりに、複数の類似した対局者で構成されるクラスに対して、その打牌選択を模倣する。続いて、提案手法の有効性を示すために行った実験について説明する。

6.1 実験環境

本研究の実験では、OSが「ubuntu 20.04」、物理CPUが1で、CPUコア数が16の「AMD Ryzen 9 3950X 16-Core Processor」を搭載した機材を用いた。

表3 模倣プレイヤーと人間プレイヤーの打牌選択の一致率

グループ平均	完全一致率
A	48.29%
B	43.70%
C	47.72%
D	48.18%
ALL	51.10%

6.2 実験内容

本研究での実験内容について、結果を伴うものは順を追って記述する。

6.2.1 使用する牌譜データについて

実験は、「天鳳」において、2012年から2018年に打たれたデータである「系列の学習データ」を、系列の出現数の一覧の作成用に用いる。2019年に打たれたデータである「系列の検証データ」を用いて模倣の精度を検証する。ある対局者の模倣が実現できたかを確認するために、プレイヤーを複数選ぶ必要があるが、これは2012年から2019年の期間中に東南戦で4000半荘以上を打っている対局者を対象とする。

6.2.2 提案手法による麻雀プレイヤーと対局者の打牌の一致率比較実験

実験では、過去にプレイヤーが最も選択した系列から求めた候補の上位5番までの牌の中に、対局者の選択した打牌が一致していれば、一致と見なしている。その結果を表3に示す。

グループA, B, C, Dは無作為にグループ化した10人のプレイヤーの平均の値を取って算出している。グループALLは対象とした人間プレイヤーと同じ卓に着座していたプレイヤーを全てグループ化し、算出している。この結果より、それぞれの個人の選択だけを学習データの系列として用いた場合よりも、全体のプレイヤーのデータを用いた方が一致率が高いという結果となり、模倣に関する優位性を示すことはできなかった。

7. 考察

事前実験より、対局者の目指す役の戦略の違いにより、系列が異なっていることが読み取れた。上位に頻出する系列についても、著者の直感と一致するものであった。色情報を無くしたことで、「染め手」と呼ばれるような同一の色の牌と字牌のみで構成される役については、系列を適用できなくなったが、何らかの形で手牌の他の色情報を知ることによって、今回検討した役以外の戦略を検討することができるのではないかと考えられる。

しかし、系列を考えるためには、人間プレイヤーの牌譜を年単位で集める必要があるため、少ない牌譜数でも対応することができるように、既存の系列の中でほとんど同じと

みなせるような遷移をまとめていく必要があると感じた。

提案手法の肝となる実験では、それぞれの個人の選択だけを学習データの系列として用いた場合よりも、全体のプレイヤーのデータを用いた方が一致率が高いという結果になった。これは、個人で対局することができる局数には限界があるため、データ数が十分ではなかったのではないかと懸念がある。しかし、平均をとったデータでは見えない点として、例えば特定の系列を用いる局面における一致率が高い可能性も否定できないため、引き続き詳しく調査を行っていく。

8. まとめと今後の予定

現在までで、系列情報による牌譜の分析を行い、対局者の打牌の傾向を探ることに成功した。特に、人間に理解できる形のみで打牌を決定し、模倣しようとした点が新しい。一方で、特定の人間プレイヤーの模倣という点では、その優位性を示すことができなかつたため、まだまだ改良の余地が残されていると感じた。加えて、実践的な対局実験を行うことができなかつたため、熟練プレイヤーを模倣した本提案手法の麻雀プレイヤーが、模倣したプレイヤーと比較してどの程度の強さを発揮するのかを調べることはできなかった。

系列の課題は、模倣相手プレイヤーの十分な数の牌譜が必要となるため、今後は系列を一般化して更に統合できる局面を探ることで、1系列が表せる手牌の形を増やし、学習データが少ない場面であっても、人間プレイヤーを模倣する方法を探りたい。また、系列は細かい手牌の形に着目しているため、人間プレイヤー毎に系列の出現数の分布を比較することで、人間の打牌選択の癖や傾向を調べることができる。これを応用して一般的な平均と比較すれば、そのプレイヤーが間違いやすい局面を可視化して、幅広いプレイヤーへの指導やプレイヤーとの協調にも活かせるのではないかと考えている。この点を踏まえ、模倣の達成度を確認するためにも「自分とどの程度似ていると感じたか」という内容のアンケート調査をいずれは検討したい。

参考文献

- [1] 水上直紀, 中張 遼太郎, 浦 晃, 三輪 誠, 鶴岡 慶雅, 近山 隆, “多人数性を分割した教師付き学習による4人麻雀プログラムの実現”, 情報処理学会論文誌, Vol. 55, pp.2410-2420, 2014.
- [2] 吉村健志, 宝珍輝尚, 野宮浩揮, “タブーサーチを用いた麻雀における最適行動の探索”, 第21回ゲームプログラミングワークショップ2016 論文集, pp. 73-80, 2016.
- [3] 塚田寛一郎, 川又泰介, 松田源立, “シケンシャルパターンマイニングを用いた麻雀の捨て牌の傾向分析”, 情報処理学会第84回全国大会講演論文集, pp.404-405, 2022.
- [4] M. Kurita and K. Hoki, “Method for Constructing Artificial Intelligence Player With Abstractions to Markov Decision Processes in Multiplayer Game of Mahjong,” in IEEE Transactions on Games, vol. 13, no. 1, pp. 99-110, March 2021, doi: 10.1109/TG.2020.3036471.