

# 麻雀初級者のための状況に応じた着手モデル選択

田中 悠<sup>†1,a)</sup> 池田 心<sup>†1,b)</sup>

**概要:** 本研究では麻雀を対象として、局面の情報から取るべき“手”ではなく取るべき“戦略”を出力する決定木の作成を行った。まず麻雀において重要な戦略である「早く上がる」「高得点狙い」「振り込み回避」といった評価基準を元に行動を評価する3つのモデルを作成、これらを用いて上級者のゲームの記録（牌譜）を上級者が意図したと思われる戦略別に分類した。上級者が意図したと思われる戦略を付与した牌譜を教師として局面の情報から取るべき戦略を出力する決定木の学習を行ったところ、決定木が元の戦略に正しく分類できた割合は約31%、決定木が導出した戦略を元にモデルを組み合わせた手の評価において上級者の手が3位以内であった確率は約86%であった。

## 1. はじめに

これまでコンピュータプレイヤーの研究の多くは強いプレイヤーを作ることとを目的とし、チェスや将棋等のボードゲームにおいては実際にトップレベルの人間プレイヤーに勝利するなど、十分な成果を上げている。現在でも、多くのゲームにおいて人間を超えることを目的とした強いコンピュータプレイヤーの研究は行われており、将来的にはほぼすべてのゲームにおいてコンピュータプレイヤーの強さが人間のトッププレイヤーを凌ぐことが予想される。そこで私は今後、人間がコンピュータプレイヤーを強くする以外に、コンピュータプレイヤーが人間を教育するというこれまでとは逆の工程の重要性が増すと考える。

人間プレイヤーを教育する際には、「このような状況だから」という理由とともに取るべき手ないし戦略を示すことが望まれるが、既存のコンピュータプレイヤーの多くは評価部分が人間には理解しづらく、手の良し悪しを示すことはできても、「なぜその手が良い手なのか?」「どのような戦略の基に導出された手なのか?」といったことは示すことができないため、教育において重要な「考えさせること」には向かない助言となってしまう。

そこで本研究では、より教育に適した助言を出力できるモデルの作成を目的として、「状況に応じて取るべき戦略を出力するモデルの作成手法」を提案する。既存のコンピュータプレイヤーの多くにおいて一種のブラックボックスとなっていた局面の評価を人間が理解しやすい評価基準を

もつ複数の単目的行動モデルによって行い、上級者のゲームの記録を上級者が選択したと思われる戦略別に分類、これを教師として局面の状況を入力とし取るべき戦略を出力する決定木を作成する。これにより、ゲームにおけるコンピュータによる初級者への教育や今後のコンピュータプレイヤーの作成に貢献する。

## 2. 麻雀の概要

本章では、本研究で対象としたゲーム、「麻雀」について説明する。

### 2.1 麻雀のルール

麻雀は麻雀牌を用いて2~4人のプレイヤーで行う多人数不完全情報ゲームである。

麻雀牌は萬子（マンズ）、筒子（ピンズ）、索子（ソーズ）の3色に一から九までの数字のついた27種類の牌（数牌）と、漢字1文字のついた7種類の牌（字牌）の計34種各4枚ずつの136枚で構成される。

ゲームの開始時に各プレイヤーに13枚の牌（手牌）が他のプレイヤーには伏せられた状態で配られ、山から牌を1枚手牌に加える（ツモ）、牌を手牌から1枚捨てる（打牌）をプレイヤーが順に繰り返して和了を目指す。和了するためには基本的に手牌に1枚加えた14枚で4つの面子（メンツ）と1つの雀頭（ジャントウ）を揃える必要があり、面子とは同種類の牌を3枚揃える刻子（コーツ）、もしくは同色の3連続の数字を揃える順子（シュンツ）の3枚1セットの牌の集合であり、雀頭とは同種類の牌2枚の牌の集合である。また、他者が打牌した牌によって面子を構成出来る場合には“鳴き”によってその牌を手牌に加えて面子を構成

<sup>†1</sup> 現在、北陸先端科学技術大学院大学  
Presently with Japan Advanced Institute of Science and Technology

a) y.tanaka@jaist.ac.jp

b) kokolo@jaist.ac.jp

したり和了することも許されている。

他プレイヤーが打牌した牌で和了った場合（ロン和了）にはそのプレイヤーのみから点数を得て、自身でツモした牌により和了った場合（ツモ和了）には他のプレイヤー全員から分割で点数を得る。和了時に得られる点数は和了時の手牌の形（役）、鳴きの有無、各局で1人設定されている親か否かなどにより決定されている。手牌の配布から、いずれかのプレイヤーが和了るか、規定回数のツモ・打牌を終えるまでを1局と呼び、規定回数局を終えるかいずれかのプレイヤーの点数が0未満となった時点での点数の多さを競う。麻雀には国や地域ごとに様々なローカルルールが存在するが、本研究ではオンライン麻雀ゲーム「天鳳」[1]で使われているルールを採用した。

麻雀は世界各国で親しまれており、多くの競技人口を有するが、「ルールが複雑であること」や「人間プレイヤーの多くが迷信や不確定な信念を持っている」といった理由により、教育用途のプログラムの必要性が高い。

## 2.2 麻雀における戦略の重要性

麻雀の目的は他プレイヤーよりも多くの点数を持って規定回数局を終えることである。他プレイヤーとの点差を広げる一番有効な手段は和了ることなのだが、4人で行う多人数ゲームであり、いずれかのプレイヤーが和了や規定回数のツモ・打牌で1局は終了となるため、1プレイヤーの局ごとの和了れる確率はせいぜい2割程度であり、和了れない時にどれだけ失点を防げるかも重要である。そのため、麻雀においては1手毎に「どの牌を打牌すれば和了やすいか」「どの牌を打牌すれば他プレイヤーにロン和了されないか（振り込まないか）」などの評価基準に基いて手の優劣を見極める技術だけではなく、その局面においてどのような戦略のもとに手を選択するのかを判断する技術も必要となる。

麻雀において初級者は往々にしてこのような戦略を判断する技術が未熟であり、例えば、「オーラス、2位とは8000点差のトップ、3位のプレイヤーがリーチをかけている中で、自身も追っかけリーチを行い、結果として3位のプレイヤーに大きい手を振り込み、3位まで順位を落とす」など、状況を考慮せずに一貫して和了を目指すため、他プレイヤーに振り込むことが多くなる、最終局面で順位に関係ない和了をする、などして結果が悪くなる。また、他のゲームの「攻め重視」「守り重視」などの戦略と比べて、麻雀では和了を完全に諦める代わりに失点を極力避ける「ベタオリ」など極端な戦略も必要であるため、適切な戦略を選択することの重要性が高い。

## 2.3 目的とする教育方針

麻雀のような1回のゲームが複数回の小ゲームで構成されるゲームにおける教育のための助言方法として以下の2つが考えられる。

1. なんらかのアルゴリズムで着手を決定し、助言としてその着手のみを示す
2. なんらかのアルゴリズムで適切な戦略を示す、あるいはそれらに基づいた着手まで示す

既存のコンピュータプレイヤーの多くは手の良し悪しを判別できてもなぜ良いのかはわからないため、教育目的に利用する際には助言方法1.の「この局面ではaを選択すべきです」「あなたが選択した手の評価値は50で、最善手のaと比較して20悪いです」といったような形式とならざるを得なかった。助言方法1.は初級者に手っ取り早く良い手を選択させるには良い手法だが、初級者に手の選択を考える余地が残されていない。

そこで、本研究では助言方法2.のような助言を出力するべく取るべき戦略を示すモデルの作成を目指す。このようなモデルを用いることにより、「この局面ではとにかく高い手を作ることを目指して手を選択すべきです」といったように、プレイヤーに手の選択の理由を教示するとともに手を考える余地を残すような助言や、「あなたが選択した手は早さでは最良ですが、この局面では失点確率の低いa1かa2がオススメです」といったような、取るべき手と併せてその理由も示す助言が可能となり、単に選択すべき手を示す助言方法よりも初級者の成長を助けると考える。

## 3. 関連研究

本章では関連研究として、麻雀を対象としてコンピュータプレイヤーを作成した研究をいくつか紹介する。

麻雀は多人数不完全情報ゲームゆえの研究の難しさ、ルールの複雑さ、様々なローカルルールの存在、などの理由により学術的な研究は少なく、一般論として麻雀のコンピュータプレイヤーはまだ人間上級者よりも弱いとされている。一方で「相手の待ち牌を捨て牌から読む」「各合法手のX巡以内の和了確率の導出」といった部分問題を解く単目的行動モデルに関しては人間の上級者よりも優秀なものが作成されている。

本研究では、比較的容易に作成でき、麻雀における有用性も認められている単目的行動モデルを麻雀を主要な戦略ごとに用意し、これを用いて牌譜の戦略別分類、取るべき手の導出などを行う。

### 3.1 ゲーム全体を行うコンピュータプレイヤー

北川らの研究[2]では3層ニューラルネットワークを用いて行動評価関数を上級者の牌譜から学習し、ゲーム全体を行う麻雀コンピュータプレイヤーを作成している。結果として、麻雀コンピュータプレイヤーが最善と判断した手と牌譜内で打たれた手との一致率はツモ局面において約56%、鳴き局面において約89%となったが、実際にインターネット麻雀にて174試合レーティング戦を行わせた結果、レートは下位13%程度のものにとどまっている。

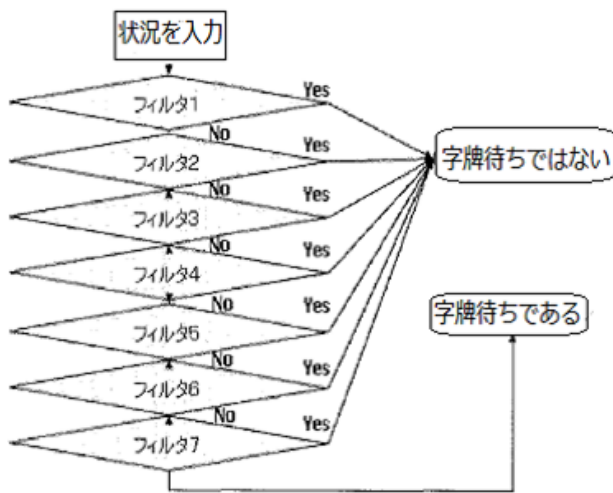


図 1 フィルタによる判別フローチャート  
Fig. 1 flowchart for decision

### 3.2 部分問題を解く単目的行動モデル

とつげき東北らの研究 [3] では、「捨て牌に占める 3~7 の数牌の比率」「シャンポン待ちが可能な字牌の数」など上級者プレイヤーの知見に基づいて設定した 7 つの判断要素を元に、リーチをかけているプレイヤーが字牌待ちであることを判別するフィルタを作成。図 1 に示すフローチャートの形でフィルタ組み合わせて「リーチをかけているプレイヤーの待ち牌が字牌か否かを判断するモデル」を作成、実験においてプロを含む 45 名の被験者と正答率を比較した結果、参加者の大部分よりも優れた結果を残している。

また麻雀においては、通常の 4 人麻雀から多人数性と不完全情報性を排除し単純化した、1 人麻雀を対象として行っている研究が多く存在する [4][5]。

### 3.3 麻雀へのクラス分類の適用

水上らの研究 [4] では 1 人麻雀と 4 人麻雀の差を解析し、その差を埋めることで 1 人麻雀コンピュータプレイヤーの 4 人麻雀への適用を図っている。その際に、4 人麻雀と 1 人麻雀の最も大きな差である“降り”を 1 人麻雀コンピュータプレイヤーに取り入れるために、降りるべき局面とそうでない局面の分類を行っている。降りるべき局面で降りる 1 人麻雀プレイヤーと常に 1 人麻雀の基準で打牌を選択するコンピュータプレイヤーに 4 人麻雀を行わせた結果、降りるべき局面で降りる 1 人麻雀プレイヤーは通常の 1 人麻雀プレイヤーと比べて、和了率こそ低くなったものの、振り込み率は下がり、平均順位も良くなっている。

## 4. 提案手法

本研究では麻雀を対象として、ゲーム中の順位、点数、手牌などの状況から取るべき戦略を出力する決定木の作成手法を提案する。

麻雀においてプレイヤーの行動は大別して牌をツモした

後に打牌を選択するかというツモ局面の行動と、他プレイヤーが捨てた牌を鳴くかどうかを選択する鳴き局面の行動の 2 種類があるが、本研究では単純化のためツモ局面の行動に絞って研究を行った。

本手法は以下の 5 手順により構成される。

1. 上級者の牌譜（局面の情報と上級者が選択した手）から上級者が打牌を選択するときの戦略を推察し列挙する
2. 1. で挙げられた各戦略ごとにそれを再現する単目的行動モデルを用意する
3. 上級者の牌譜の合法手それぞれを、単目的行動モデルで評価を行って順位づけし、実際の打牌がどのタイプだったのかをラベル付けする
4. ラベル付きの牌譜を教師データとして局面の情報からタイプを予測する決定木を作成、タイプの正答率を評価する
5. 4. で作成した決定木により、局面の入力から取るべき戦略を出力する教育プログラムを実現する

## 5. 単目的行動モデルの設計と評価

本章では、本研究で用いた単目的行動モデルの選定方法、設計手法およびに評価実験の結果を示す。

### 5.1 上級者がもつ戦略の読み取り

まず、上級者が持つ戦略を列挙するために「天鳳」の麻雀大会、天鳳名人戦の決勝トーナメントの 2 半荘分の打牌、計 991 手から手動で戦略を推察した。その結果、上級者の全ての打牌は以下の 5 つの戦略の組み合わせにより説明することができた。

#### 早い和了を目指す

和了を早く達成できそうな打牌選択。上級者の打牌選択時の戦略の大半を占めており、1 局の開始直後に特に多く見受けられた。

#### 高得点を目指す

和了時に高い得点を見込める打牌選択。順位が 3 位以下であるときなどに多く見受けられた。

#### 振り込みを避ける

他のプレイヤーへの振り込みを極力避ける打牌選択。他プレイヤーがリーチをかけている時、他プレイヤーの聴牌（和了に必要な牌が残り 1 枚となった状態）が濃厚な局の終盤、振り込みによって順位が下がることを避けたいゲームの終盤などによく見受けられた。

#### 安牌（振り込む危険性の低い牌）を残しておく

後々、他プレイヤーがリーチをかけた時や他プレイヤーの聴牌が濃厚となった時のために安牌を手牌に残しておくための打牌選択。

#### 和了やすい形をとる

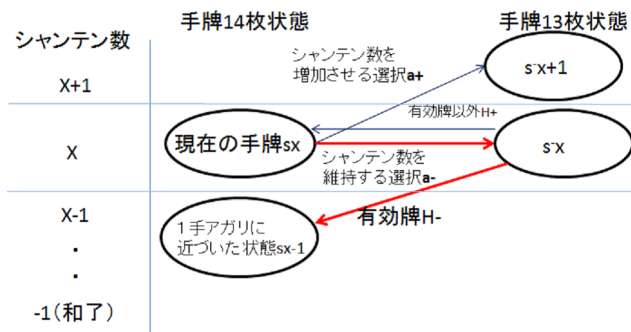


図 2 打牌とツモによるシャンテン数の状態遷移図

他プレイヤーが切りそうな牌で待てるようにしたり、聴牌時の待ちの形をよくしようという打牌選択。

本研究では、これらの戦略全てを単目的行動モデル化するのではなく、戦略に沿った単目的行動モデルの実装しやすさ、戦略の登場頻度、初級者にとっての理解しやすさなどから、「1. 早い和了を目指す」「2. 振り込みを避ける」「3. 高得点を目指す」の3つの戦略の単目的行動モデルを作成し、上級者の戦略の読み取りを試みた。

## 5.2 単目的行動モデルの設計

本節では3つの単目的行動モデルの設計を述べる。

### 5.2.1 「早い和了を目指す」単目的行動モデルの設計

当モデルは、自分の手牌と山に残っていると思われる牌（見えていない牌）の種類と枚数という入力から、各打牌選択時の「和了までの早さ」の導出を目的として設計した。

麻雀において早い和了を目指す際には、打牌による手牌のシャンテン数（手牌が聴牌となるまで最小で何回牌を入れ替える必要があるかを示す値）の変化と、有効牌（手牌に加える事でシャンテン数を下げることのできる牌）の残り枚数が重要となる。図2はプレイヤーが牌をツモし打牌するまでの状態を手牌14枚状態 $s$ 、打牌した後の状態を手牌13枚状態 $s^-$ とした時の打牌とツモによる状態遷移を示す。

シャンテン数 $x$ の手牌14枚状態 $s_x$ において全ての打牌は、打牌後シャンテン数 $x$ の手牌13枚状態 $s_x^-$ となる打牌 $a^-$ か、打牌後シャンテン数 $x+1$ の手牌13枚状態 $s_{x+1}^-$ となる打牌 $a^+$ のいずれかに分類される。このとき、 $a^-$ が和了に早める打牌選択であり、 $a^+$ は和了を遅くする打牌選択であるといえる。和了状態（シャンテン数が-1）以外の手牌14枚状態において $a^-$ は確実に存在し、 $a^+$ と比較して和了が遅くなるということが稀であるため、 $a^+$ の各打牌の評価値は-（まだ見えていない対象の牌の数）という計算量が少なく簡単な値に設定した。一方で、 $a^-$ の評価値は各打牌毎に正確に比較する必要が有るため、 $a^-$ を行った後の手牌13枚状態での有効牌の枚数 $H^-$ とするように設定した。これにより、評価値の高い打牌を選択していくことで早い和了を期待できる。

### 5.2.2 「高得点を目指す」単目的行動モデルの設計

当モデルは、自分の手牌、山に残っていると思われる牌（見えていない牌）の種類と枚数、ドラという入力から、各打牌選択時の「和了時の得点の高さ」の導出を目的とし設計した。

麻雀において、高得点で和了するには難易度が高いがそれ一つで高得点となる役を成立させるか、難易度が比較的低い役を多く成立させることが鍵となる。本研究では「高得点を目指す」単目的行動モデルとして、数ある麻雀の役の中でも特に出現頻度が高く、他の様々な役と重複しやすいドラ（手牌に含めて和了することで得点が高くなる牌）とタンヤオ（2~8の数牌のみでの和了形に付く役）を多く和了に絡めるモデルを作成した。

当モデルでは和了時の得点を高めるために、ドラとタンヤオに必要な牌を手牌に残しやすくするよう手の評価を行うほかに、「早い和了を目指す」単目的行動モデルでも用いた打牌によるシャンテン数の変化の概念を取り入れ、シャンテン数 $x$ の手牌14枚状態からシャンテン数 $x+1$ の手牌13枚状態になるような打牌 $a^+$ も手牌に残しやすくするよう設定した。これは、1順ごとに1枚牌を入れ替えることができ、他プレイヤーの和了等がなければ1局につき20回程度ツモ順が回ってくるという麻雀の性質上、どのような役を目指しても局の序盤では実現確率は基本的に0にはならず、和了達成時の得点の高さのみ手の評価を行うと、ほぼ和了することのできないモデルになってしまうためである。これにより、評価値の高い打牌を選択していくことでそこそこ和了確率を持ち、「早い和了を目指す」単目的行動モデルよりも和了時の平均得点が高いモデルとなることが期待できる。

### 5.2.3 「振り込みを避ける」単目的行動モデルの設計

当モデルは、他プレイヤーの捨て牌、順目、リーチの有無という入力から各打牌選択時の「安全さ（振り込みとならない確率）」の導出を目的とし設計した。安全さの導出のため本モデルでは、全合法手Aについて打牌 $a$ がいずれの他プレイヤーにも振込とならない確率を、表1のプレイヤーが聴牌確率していた場合の牌種 $kind(a)$ ごとの待ち牌である確率 $p_{kind}(kind(a))$  [6]と、表2の上級者の牌譜35半荘分のデータから統計をとった順目 $t$ におけるプレイヤー $i$ の聴牌確率 $p_{turn}(t_i)$ を用いて、式1により導出する。

これにより、評価値の高い打牌を選択することで他プレイヤーへの振り込みを避けることの可能性の高いモデルとなることが期待できる。

$$p_{safe}(a) = \prod_{i=1}^3 (1 - p_{kind}(kind(a)) \times p_{turn}(t_i)) \quad (1)$$

## 5.3 単目的行動モデルの評価

本節では作成した単目的行動モデルの評価実験の結果を

表 1 牌種ごとの聴牌しているプレイヤーへの振り込み危険度

Table 1 Dangerousness of each tile

牌種 (k)	待ち牌である確率 $p_{kind}(k)$ (%)
無スジ 4,5,6	12.3
無スジ 3,7	7.1
片スジ 4,5,6	7.0
無スジ 2,8	7.0
無スジ 1,9	6.3
スジ 3,7	5.5
両スジ 4,5,6	4.8
スジ 2,8	4.8
単騎以外で待てる字牌	3.4
スジ 1,9	2.9
単騎でのみ待てる字牌	0.9

表 2 順目毎の聴牌確率

Table 2 Probability of ready hand

順目 (t)	聴牌数 / サンプル数	聴牌確率 $p_{turn}(t)$ (%)
1	1 / 4612	0.02
2	6 / 4600	0.13
3	31 / 4598	0.67
4	64 / 4573	1.40
5	128 / 4507	2.84
6	232 / 4385	5.29
7	344 / 4195	8.20
8	422 / 3912	10.79
9	481 / 3544	13.57
10	494 / 3169	15.59
11	480 / 2780	17.27
12	476 / 2391	19.91
13	421 / 2017	20.87
14	351 / 1654	21.22
15	289 / 1381	20.93
16	226 / 1133	19.95
17	183 / 934	19.59
18	151 / 631	23.93
19	56 / 166	33.73
20	10/20	50.00
リーチ中	—	100.00

示す。

### 5.3.1 「早い和了を目指す」単目的行動モデルの評価

当モデルは、通常の麻雀ではなく 1 人麻雀にて評価を行う。1 人麻雀とは、手牌を与えられた後、ツモと打牌のみを繰り返して和了を目指す、単純化された麻雀である。通常の麻雀と比較すると多人数性や不完全情報性が排除され、鳴き局面も存在しないが、本研究ではツモ局面の行動のみを対象としていること、通常の麻雀で強いプレイヤーほど 1 人麻雀での和了り率も高いこと [7] から、本モデルの評価に使用した。

評価においては、同様に 1 人麻雀の和了率を検証に使用した水上直紀らの研究 [4] と同じく、27 回のツモ以内の和

表 3 1 人麻雀における和了率

Table 3 Probability of winning in single player Mahjong

プレイヤー	和了率*1 (%)
人間上級者プレイヤー	51
関連研究の 1 人麻雀モデル	48
本研究のモデル	45.9
人間平均プレイヤー	36

表 4 1 人麻雀における和了率、和了時得点、役構成

Table 4 Details of completed hand in single player Mahjong

	和了時			
	和了率 (%)	和了時 平均得点	タンヤオ率 (%)	和了時 平均ドラ数
高得点モデル	30.1	6214	71.4	0.76
早和了モデル	45.9	3586	9.2	0.41

了率により行い、同研究のモデルや人間プレイヤーと比較することで検証した。プレイ数は本研究のモデルのみ 1000 回で他は 100 回、人間上級者プレイヤーとは「天鳳」において最上位の鳳凰卓でプレイできるプレイヤーであり、人間平均プレイヤーは上位 50%ほどのプレイヤーである。表 3 に実験の結果を示す。

本モデルの和了率は、上級者や関連研究の 1 人麻雀モデルの和了率を越えはしなかったが、平均プレイヤーよりも上級者プレイヤーに近い和了率を記録しており、早い和了を目指すという目的を十分に満たしている。

### 5.3.2 「高得点を目指す」単目的行動モデルの評価

当モデルは、「早い和了を目指す」単目的行動モデルの検証と同様に、1 人麻雀により検証を行った。「早い和了を目指す」単目的行動モデルと和了率の他に和了時の得点と「高得点を目指す」単目的行動モデルにおいて和了時に多く絡めるよう調整したタンヤオとドラの両役出現率を比較した。プレイアウト数は 1000 回、ドラは 1 局につき 1 枚、得点はどちらも親の時の得点を使用した。表 4 に実験の結果を示す。

「高得点を目指す」単目的行動モデルにおいて多く和了に絡めるよう調整したドラとタンヤオの出現率は、「早い和了を目指す」単目的行動モデルと比較して目に見えて増加し、和了時の得点も約 1.7 倍となっている。和了り率に関しては、下がってしまっているが許容範囲であり、そこそこの和了確率を持ちながら、高い得点を目指すという目的に沿ったモデルを作成できた。ここから更に和了時平均得点を高めるには、他の役も考慮する、評価値の微調整を行うなどの方法が考えられるが、基本的に和了時平均得点と和了率はトレードオフの関係にあり、現状の和了率と和了時平均得点でも十分目的を達成できると考える。

### 5.3.3 「振り込みを避ける」単目的行動モデルの評価

「振り込みを避ける」単目的行動モデルの検証では、「天

\*1 有効数字が異なるのはプレイ数が既存文献では少ないためである。

表 5 振込の予測精度の検証

Table 5 Accuracy of estimation of dangerousness

予測確率 (%)	0~2	2~4	4~6	6~8	8~10	10~15	15~
実測確率 (%)	1.3	3.5	7.5	8.0	11.7	13.2	18.2
予測回数	533231	111096	29438	25001	26596	13833	4132
実測回数	6870	3861	2211	1997	3101	1821	754

鳳」の牌譜からランダムに選んだ 31 半荘, 22828 のツモ局面において全牌種について 1 人以上のプレイヤーに振り込む確率を予想 (予測確率) させ, 実際に 1 人以上のプレイヤーがその牌で待っていた割合 (実測確率) を予測確率別に集計した. 表 5 に実験の結果を示す.

予測確率が 0% となっていながらその牌で待っているとされているものは, 相手がフリテン (他者の打牌で和了ることのできない状態) の局面である. 実測値が予測の範囲を外れてしまっている結果もあるが, 概ね予測通りであり, 予測した値が高ければ実際に当たる確率も高くなっている.

単目的行動モデルによる局面のクラス分類の際には, 各モデルの打牌に対する評価値ではなく, 合法手中の評価順位を用いるため, 本モデルは本研究に用いるのに十分な性能を有していると言える.

## 6. 決定木の学習

本章では, 決定木の学習データとするための牌譜の戦略別分類法, 決定木の学習方法, 決定木の学習結果を記す.

### 6.1 学習データの用意

決定木の学習データを得るため, 牌譜を 1 局面ずつ, 上級者の打牌の前章で示した 3 つの単行動目的モデルによる合法手中の評価順位とタイプごとに設定された重みを用いて求められる  $\text{rank}(a) (=w_H \cdot \text{rankH}(a) + w_T \cdot \text{rankT}(a) + w_A \cdot \text{rankA}(a))$  を基に, 図 3 の要領で上級者が意図したと思われる戦略別に 8 タイプに分類した. 図に示す分類割合は上級者の牌譜 122 半荘分, 61771 局面を分類した際の結果である.

タイプ 1 が上級者の手が全ての単目的行動モデルで高く評価されている局面, タイプ 2~4 が 2 つの単目的行動モデルで高く評価された局面, タイプ 5~7 が 1 つの単目的行動モデルでのみ高く評価された局面である. タイプ 1~7 に分類されなかった局面に関しては, 決定木の学習の際には取り除く. このようなタイプ不明の局面が約 20% ある理由としては, 上級者の牌譜から見て取れた戦略全てを単目的行動モデルとして分類に使用しなかったことや, 高得点を目指すモデルにおいてタンヤオとドラ以外の役を考慮していないためであると推測される.

### 6.2 学習方法

表 6 が本研究で決定木の学習に使用した特徴量の一覧で

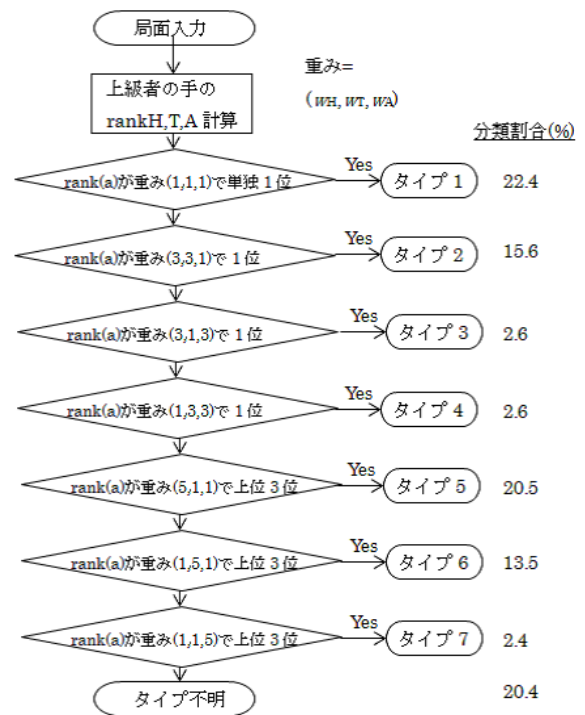


図 3 タイプ判別フローチャート

Fig. 3 flowchart for classification

表 6 決定木の学習に使用した特徴量一覧

Table 6 Features for educating decision tree

特徴量	取りうる値
自身の順位	離散値 {1,2,3,4}
局数	連続値 (正の整数)
サドンデス	離散値 {TRUE,FALSE}
現在の順目	連続値 (正の整数)
自身のシャンテン数	離散値 {-1,0,1,2,3,4,5,6}
他プレイヤーのリーチ数	離散値 {0,1,2,3}
自身の鳴きの数	離散値 {0,1,2,3,4}
1 位との点差	連続値 (整数)
1 位との点差/残り局数	連続値 (整数)
持ち点 2000 点以下のプレイヤーの有無	離散値 {TURE,FALSE}

ある. これらの特徴量から前節で述べた 7 つのタイプを予測する決定木を, データマイニングツール Weka において, 決定木生成のアルゴリズム C4.5 を基に実装された J4.8 と呼ばれるアルゴリズムを用いて作成した.

### 6.3 学習結果

本節では, 学習した決定木の性能を評価した結果を述べる.

前節で示したデータの内, 不明と分類されたものを除いた, 49179 件の局面のデータより J4.8 アルゴリズムにより決定木を学習し, 交差検証法により入力された局面のタイプを正しく導き出せた割合 (正答率) により検証する. 入力された局面のタイプごとの正答率を表 7 に示す.

決定木により入力された局面のタイプを正しく導き出せ

表 7 入力タイプ別の正答率

Table 7 Accuracy of classification for each type

タイプ	T1	T2	T3	T4	T5	T6	T7
正答率 (%)	47.6	41.9	4.2	3.1	28.7	19.7	4.9

た割合は 32.8%であった。この値は、全ての局面をランダムにタイプを決定した場合や、一番割合の多い T1 へと分類した場合よりは高いものの、うまく機能しているとは言いきれない値である。

入力タイプ別の正答率では特に T3, T4, T7 が低く、これらはいずれも「振り込みを避ける」単目的行動モデルの評価順位を重視したタイプである。そのため、上級者の打牌が「振り込みを避ける」単目的行動モデルに高く評価されている局面を精査し、関連性が高いと思われる特徴量の追加や正答率を下げる要因となるような不要な特徴量の削除を行うことで、正答率の高められると推測する。

## 7. 評価実験

4章で述べた手順に基づき、ここまで (1) 上級者の戦略の列挙, (2) 各戦略の再現のための単目的モデル作成, (3) このモデルを持ちいた上級者の打牌の戦略タイプ分け, (4) 局面から戦略タイプを導く決定木の学習, を行ってきた。本章ではこれらを用い, (5) 実際にある局面で予想した取るべき戦略から打つべき牌の決定, を行い, 評価する。

### 7.1 上級者の打牌予測

本節では決定木が予想したタイプを基に単目的行動モデルを組み合わせ, 上級者の打牌を再現できるか検証した結果を述べる。各合法手の評価値を, 前章で rank(a) を求めるのに使用した式の重みに決定木が予想したタイプの重みを使用することで求め, 上級者の打牌の評価値が合法手内で 3 位以内であった場合に, 上級者の打牌を予測できたものとし, 予測できた割合により評価を行った。表 8 に予想したタイプの正否と手の予測精度の相関を示し, 表 9 にそれらを更に入力のタイプと上級者の手の順位で細分化したものを示す。

評価の結果, タイプが一致していれば全ての局面において上級者の打牌を全合法手の中で上位 3 位以内と予測でき, タイプが一致していなくても約 80% の確率で予測できている。決定木のタイプ正答率は約 30% であるため, 全体としては上級者の打牌を上位 3 位以内と予測できる確率は約 86% となる。

合法手の数は鳴きの有無や手牌内での牌種の重複数に左右されるが, 平均 10 以上の合法手がある中で, 一般的なプレイヤーが得られるゲーム中の情報のみから, この精度で上級者の打牌を全合法手の中から 3 位以内と予測できるのは良い結果だといえる。

一方で, タイプごとに元のタイプと予想したタイプが不

表 8 タイプの正答率と予測精度

Table 8 Accuracy of estimating advanced player's hand

	割合 (%)	3 位以内率 (%)
タイプ一致	31.1	100
タイプ不一致	68.9	80.5

表 9 タイプ別上級者の打牌予測精度

Table 9 Accuracy of estimation for each type

元のタイプ	決定木の予測正誤	上級者の手の予想順位			
		1 位 (%)	2 位 (%)	3 位 (%)	3 位外 (%)
タイプ 1	一致	100	0	0	0
	不一致	94.6	4.1	0.1	1.2
タイプ 2	一致	100	0	0	0
	不一致	85.9	8.5	4.1	1.5
タイプ 3	一致	100	0	0	0
	不一致	35.8	37.4	16.8	10.0
タイプ 4	一致	100	0	0	0
	不一致	22.7	37.6	19.4	20.2
タイプ 5	一致	3.9	55.9	40.2	0
	不一致	0.6	62.7	23.2	13.5
タイプ 6	一致	4.2	48.1	47.6	0
	不一致	0.1	23.2	26.1	50.6
タイプ 7	一致	16.7	50	33.3	0
	不一致	0	1.8	8.7	89.4

一致であった場合の打牌の予測精度には大きな差があり, T5, T6, T7 の「ある戦略のみを重視」とされるタイプにおいては不一致率は特に高くなっている。このようなバラツキの原因としては, 各単行動目的モデルが出力する評価値の合法手内での重複率の違い, 6.1 節で記した各タイプへの分類条件の違いなどが一因となっていると考えられる。

今後は, 予測精度の向上や, タイプによる予測精度の差の縮小といった改善が必要とされる。

### 7.2 本モデルの有効性

上級者の牌譜の様々な特徴量から学習した決定木により, 局面毎に取るべき戦略を導出することができるようになった。単目的行動モデルのみではどのような局面においてもバランス重視の手しか出力できなかったが, 決定木によって出力された戦略ごとに適切な単目的行動モデルの組み合わせ時の重みを設定することによって, 麻雀において重要な要素である戦略を考慮した手を出力できるようになった。

例えば, 「2 着と一万点以上の差を付けてのトップで迎えたオーラス, 他プレイヤー全員がリーチを掛けているという局面において, 自身も和了れば高得点を期待できる手を聴牌したが, 聴牌を維持するには振り込む可能性の高い牌を打牌する必要があり, 安牌を切ると聴牌が崩れてしまう」という局面において, 上級者であれば一般的に, トップを維持するために聴牌を崩し安牌を切るが, 3 つの単目的行動モデル (「早い和了を目指す」「高得点を目指す」「振り

込みを避ける) のバランス重視で手を選択する場合には、「早い和了を目指す」「高得点を目指す」という2つの観点から推奨される聴牌を維持し振り込む可能性の高い牌を選択してしまう。一方で、単行動目的モデルに加えて、本稿で提案した取るべき戦略を出力する決定木を用いることで、この局面は降りるべきであると判断でき、「振り込みを避ける」単目的行動モデルの評価を重視することによって安牌を切る選択を取ることができる。

## 8. まとめ

本稿では、既存のコンピュータプレイヤーでは難しかった、局面から取るべき戦略を導出するモデルの作成手法を提案し、3つの単目的行動モデルの評価に基づいて局面を7つの戦略タイプに分類し、教師あり学習によって作成した決定木が予想した戦略タイプの正答率を測る「戦略の正答率」と、決定木が予想したタイプと局面の情報から上級者の手を予想する「手の正答率」により有効性を確認した。「戦略の正答率」についてはあまり良い結果を得ることはできなかったが、決定木の学習データに特徴量の追加や不要な特徴量の削除を行うことにより「戦略の正答率」「手の正答率」共に改善が期待できる。

正答率の改善以外の課題として、今後は当手法で作成したモデルをより教育に適したものとするために、局面から導出した戦略や手を初級者にも理解しやすい助言として自動で出力させることが挙げられる。初級者の教育効率などを指標として助言の改善を重ねることにより、コンピュータによる人間への新しい教育スタイルの確立にも期待できる。

## 参考文献

- [1] オンライン対戦麻雀ゲーム 天鳳: <http://tenhou.net/> (2014).
- [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, 情報処理学会シンポジウム論文集, No. 12, pp. 76-83 (2007).
- [3] とつげき東北, 伊藤毅志: 牌譜の解析による麻雀の分析, 人工知能学会誌, Vol. 24, No. 3, pp. 355-360 (2009).
- [4] 直紀水上, 遼太郎中張, 晃 浦, 誠 三輪, 慶雅鶴岡, 隆 近山: 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, ゲームプログラミングワークショップ 2013 論文集, pp. 1-7 (2013).
- [5] 遼太郎中張, 直紀水上, 晃 浦, 誠 三輪, 慶雅鶴岡, 隆 近山: LinUCBの1人麻雀への適用, ゲームプログラミングワークショップ 2013 論文集, pp. 114-117 (2013).
- [6] とつげき東北: 科学する麻雀, 講談社現代新書 (2004).
- [7] とつげき東北: おしえて! 科学する麻雀, 洋泉社 (2009).