

Title	初級者の教育を目的とした状況に応じた着手モデル選択
Author(s)	田中, 悠
Citation	
Issue Date	2014-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/12043
Rights	
Description	Supervisor:池田 心, 情報科学研究科, 修士

修 士 論 文

初級者の教育を目的とした
状況に応じた着手モデル選択

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

田 中 悠

2014年3月

修 士 論 文

初級者の教育を目的とした 状況に応じた着手モデル選択

指導教員 池田 心 准教授

審査委員主査 池田 心 准教授
審査委員 飯田 弘之 教授
審査委員 白井 清昭 准教授

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

1110201 田中 悠

提出年月: 2014年2月

概要

ゲームのコンピュータプレイヤは、コンピュータそのものの性能向上と人工知能技術の発展に伴い多くのゲームでその“強さ”を向上させてきた。チェスでは1997年にカスパロフが敗れ、将棋でも近年相次いでプロ棋士が敗れ、囲碁ですらモンテカルロ木探索の登場によってアマチュア高段者レベルに達するなど、その強さの向上は目覚ましく、多くのプレイヤにとっては十分なレベルにすら到達しつつある。

これら古典的なボードゲームに止まらず、より複雑な特徴を持ったゲームでのコンピュータプレイヤのアルゴリズムの研究も行われるようになってきている。例えば確率的状態遷移をする「ぷよぷよ」、不完全情報性を持つポーカー、多人数での関係性が重要になる「三国志」や「大戦略」、そしてこれらを合わせ持つStarcraftや麻雀などは、困難な対象として注目を集めている。

ゲームごとに、用いられる手法もさまざまである。チェスや将棋では従来は状態評価関数を人間が定義して、 $\alpha\beta$ 法などで木探索をするのが基本的なアルゴリズム開発の接近法であった。エキスパートシステムによってif-thenルールで記述されたコンピュータプレイヤもしばしば作られた。最近では、機械学習や最適化手法を使って上級者の棋譜から評価関数を学習することがしばしば行われる。ボナンザ法によって状態評価関数を棋譜から学習したり、Bradley-Terry法など行動評価関数を学習してさまざまに利用する方法、その他ニューラルネットワーク(ANN)やサポートベクターマシン(SVM)を用いることも多い。他方探索でもモンテカルロ木探索など、新しい方法が多く開発されている。

このように複雑で高度な技術によって、コンピュータプレイヤの強さは向上したが、その行動の理由を推測することはより困難になったと言える。例えば、状態行動型のニューラルネットワークが「着手Aが最適」という答えを返してくれたとしても、それがなぜ最適なのかは分からず、例えば初級者がゲームの考え方を理解するには役立たない。先述したように昨今コンピュータゲームプレイヤの強さは多くの人間にとって十分なものになりつつあり、今後は「面白さ」「教育性」などを重視した研究が盛んになっていくと考えている。そこで本研究では、麻雀を対象として、初級者中級者にとって理解が少しでも容易になるように情報提供を行うための仕組みを考える。

麻雀は日本では非常に人気のあるゲームであり、特に一定の年齢層以上ではルールを全く知らない人は殆どいない程のゲームである。囲碁や将棋と異なり初級者レベルでも“あがる”ことができるため、多少技量差があっても真剣勝負はともかく友達同士・同僚・家族などで遊ぶには困らないという優れた特徴を持つ。一方、趣味の多様化・核家族化・若者の社会的負荷増大などに伴い、4人集まって麻雀を遊んだり教えたりするような機会が減ってきているのも事実である。インターネット上の麻雀サイトで遊ぶことはできても、直接教えてもらえる機会は少ない。

麻雀では、“役”を覚えたり点数を計算したり和了りやすい手にすることは勿論であるが、「状況に応じて戦略を変える」ことがとても重要である。ある状況ではとにかく早く確

実に和了り，ある状況ではできるだけ高い手で逆転を狙い，ある状況では和了りを諦め，“振込み”のリスクを最小限に抑える必要がある．これは十数局の合算で順位が決まり，またランダム性のある麻雀ならではと言えるかもしれない．しばしば初級者はこの戦略判断を誤り，「あがればそれで嬉しい」といった行動原理に基づいて着手を選択し，結果として負けを早めたり，場を白けさせてしまう．

本研究ではこの戦略選択に着目し，教育目的に利用することを考えた．まず，上級者の牌譜（ゲーム記録）を入念に調査し，殆どの着手が5つの戦略によって説明できることを発見した．そこでその中から特に重要な「早い和了を目指す」「高得点を目指す」「振り込みを避ける」の3つの戦略に注目し，まずこれらのみを考えて着手を評価する“単目的行動モデル”を作成した．一人麻雀による評価実験では，例え早い和了を目指すモデルは，高得点を目指すモデルに比べて1.5倍の頻度であがれるが，逆に高得点を目指すモデルは平均して1.7倍の点数で和了ることができることが確認できた．

その上で，上級者が「どの状況で，どの戦略を使っているか」をラベル付けし，さらにそれを決定木による機械学習手法であるJ4.8を用いて学習させた．ラベルは「全ての戦略を均等に重視」「ある戦略のみを重視」「2つの戦略を重視」の全7通りを用いた．この決定木により，手牌などの状況を入力すると，「どの戦略を重視すべきか」「だからどの手が好ましいか」などが出力されるシステムを作成することができた．上級者の着手をシステムが3位以内とした率（汎化性能：3位以内率）は86%に達した．

これにより，従来ならば「最善手は5ピンです」「最善手は5ピンで，勝率は34.8%です」あるいは「最善手は5ピンです，次善手は1ピンです，評価値はそれぞれ54と47です」といった情報しか提示できなかったものが，「この局面ではとにかく安全さを重視しましょう．最善手は5ピンで振込む確率は1.5%，1ピンだと2.1%です」あるいは「この局面は高さで安全さを重視しましょう．1ピンなら早く，6ピンならより高いのですが1ピンは安く，6ピンは振込む危険が6.2%とやや大きいです．5ピンはバランスが取れています」といった教育が可能になる．

本研究により，麻雀人口の増大，プレイヤーの不満や負担の軽減のみならず，他の戦略選択が重要なゲームへの応用も進むことを期待したい．

目次

第1章	はじめに	1
1.1	背景	1
1.2	論文の構成	1
第2章	対象とするゲーム	3
2.1	麻雀のルール	3
2.2	麻雀における戦略の重要性	4
2.3	目的とする教育方針	4
第3章	関連研究	6
3.1	ゲーム全体を行うコンピュータプレイヤー	6
3.2	部分問題を解く単目的行動モデル	6
3.3	麻雀へのクラス分類の適用	7
第4章	提案手法の概要	9
第5章	単目的行動モデルの設計と評価	10
5.1	上級者がもつ戦略の読み取り	10
5.2	単目的行動モデルの設計	11
5.2.1	「早い和了を目指す」単目的行動モデルの設計	11
5.2.2	「高得点を目指す」単目的行動モデルの設計	12
5.2.3	「振り込みを避ける」単目的行動モデルの設計	13
5.3	単目的行動モデルの評価	15
5.3.1	「早い和了を目指す」単目的行動モデルの評価	15
5.3.2	「高得点を目指す」単目的行動モデルの評価	15
5.3.3	「振り込みを避ける」単目的行動モデルの評価	16
第6章	決定木の学習	17
6.1	学習データの用意	17
6.2	学習方法	18
6.3	学習結果	19

第7章	評価実験	21
7.1	上級者の打牌予測	21
7.2	本モデルの有効性	23
第8章	まとめ	24
	謝辞	26

第1章 はじめに

1.1 背景

本論文ではゲームにおいて、局面を把握し着手を選択するプログラムをコンピュータプレイヤーと呼称し、ゲーム全体ではなくゲーム内の特定の目的を達成するための、状況を入力とし全合法手の評価値を出力する入出力モデルを単目的行動モデルと呼称する。

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が目的とし、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ふ。

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。

1.2 論文の構成

以降の本論文の構成は以下のとおりである。

第2章 対象とするゲーム

本研究が対象とするゲーム、「麻雀」のルールや人間プレイヤーを教育する上での問題点などを述べる。

第3章 関連研究

着手モデルの選択に用いる決定木の作成アルゴリズムや、既存の麻雀コンピュータプレイヤーを作成した研究、上級者レベルの単目的行動モデルを作成した研究を紹介する。

第4章 提案手法の概要

本論文が提案する手法の全体像を述べる。

第5章 単目的行動モデルの設計と評価

本研究で使用した3つの単目的行動モデルについて、それらの選定方法、設計方法、評価結果を述べる。

第6章 決定木の学習

決定木の学習に用いるデータの概要や、学習手法、評価実験の結果を述べる。

第7章 評価実験

決定木が予想したタイプを元に3つの単目的行動モデルを組み合わせて得た手と上級者の着手との一致率により手法の有効性を確認する。

第8章 まとめ

本論文のまとめと、今後の展望や課題を述べる。

第2章 対象とするゲーム

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

先述の通り、麻雀は複数回の局を終えた時点での点数の多さを競うゲームである。他プレイヤーよりも多くの点数を持って規定回数の局を終えることが目的である。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1つの評価基準の元に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、適切な戦略を取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、助言方法1.の「この局面ではaを選択するべきです」「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」といったような助言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おらず、教育の観点からはあまり良くない。

そこで、本研究では助言方法2.の取るべき戦略を示すモデルの作成を目指す。この方法ならば、「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るべきです」「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

を残すことができ、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。

第3章 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。2.3節で示したような教育法に関する研究は、適当なものを見つけ出すことができなかった。

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。

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、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、表3.1のような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、これらを図3.1に示すフローチャートの形で組み合わせて「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].

表 3.1: 字牌待ち判別のための7種類の評価基準

	判断要素	字牌待では無いと判断する基準
フィルタ 1	字牌の一般的な危険性	当該指標が 0.25 を下回れば Yes
フィルタ 2	捨て牌に占める 3~7 の数牌の比率	40% 未満であれば Yes
フィルタ 3	序巡好牌切りの数	初めて 2~8 牌を切るまでに, それ以外の牌を 3 種類以上切っている
フィルタ 4	残り無スジの数	計 18 本のスジのうちフリテンにならない本数が 13 本以上
フィルタ 5	シャンポン待ちが可能な字牌の数	場に 3 枚以上見えている字牌が 5 種類以上
フィルタ 6	シャンポン待ちが可能な字牌の数	場に 2 枚以上見えている字牌が 6 種類以上
フィルタ 7	生牌字牌が少ないか	場に 1 枚も見えていない字牌が 0 種類以上

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。

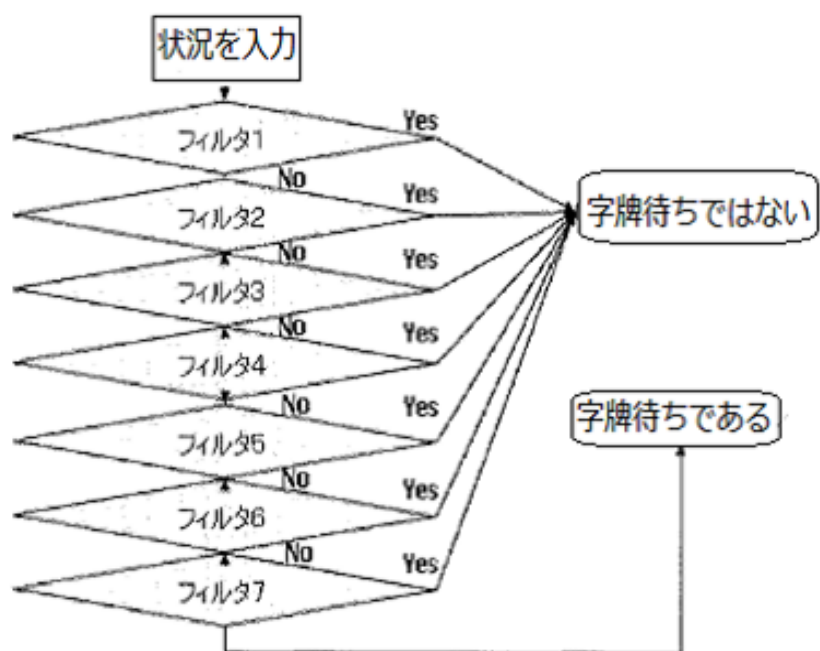


図 3.1: フィルタによる判別フローチャート

第4章 提案手法の概要

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、本研究では簡単化のためツモ局面の行動に的を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

第5章 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

まず上級者がどのような戦略のもとに手を選択しているのかを「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手から自身の知見に基づいて推察した。

各戦略については推察前に用意するのではなく、推察時に新しい戦略と推察した際に逐次追加していった。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

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。

上級者の打牌から見受けられたこれらの5つの戦略の内、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（手牌に加える事すること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5.1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あるため、 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。

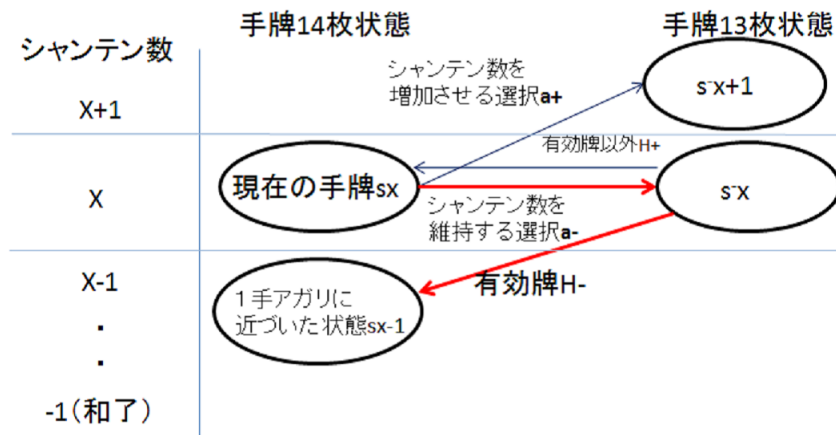


図 5.1: 打牌とツモによるシャンテン数の状態遷移図

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、「ドラ以外の牌の評価値を高くする」「タンヤオに不要な牌（1,9,字牌）の枚数の少なさに応じてタンヤオに不要な牌の評価値を高くする」という処理を行っている。また麻雀では、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、和了時の得点の高さのみで評価を出力するとほぼ和了することのできないモデルになってしまう。そのため当モデルにおいても、「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^+ には評価値にはマイナスを課すよう設定した。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 5.1[6] 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 $p_{kind}(kind(a))$ と、表 5.2 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 5.1 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。

$$p_{safe}(a) = \prod_{i=1}^3 (1 - p_{kind}(kind(a)) \cdot p_{turn}(t_i)) \quad (5.1)$$

表 5.1: 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

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 5.2: 順目毎の聴牌確率

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

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5.3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せている。

表 5.3: 1人麻雀における和了率

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

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5.4に実験の結果を示す。

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。

「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。

表 5.4: 1人麻雀における和了率, 和了時得点, 役構成

プレイヤー	和了率 (%)	和了時 平均得点	和了時 タンヤオ率 (%)	和了時 平均ドラ数
「高得点を目指す」	30.1	6214	71.4	0.76
「早い和了を目指す」	45.9	3586	9.2	0.41

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 5.5 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。

表 5.5: 振込の予測精度の検証

予測確率 (%)	0	0~2	2~4	4~6	6~8	8~10	10~15	15~20	20~
実測確率 (%)	3.0	1.3	3.5	7.5	8.0	11.7	13.2	17.4	25.1
予測回数	32825	533231	111096	29438	25001	26596	13833	3697	435
実測回数	980	6870	3861	2211	1997	3101	1821	645	109

第6章 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、現在の順位、自身の手牌、他プレイヤーの捨て牌など、ゲーム中においてプレイヤーの知ることのできる全ての情報と上級者の打牌が記された牌譜を1局面ずつ、上級者の打牌の前章で示した「早い和了を目指す」単目的行動モデル（早さ）、「高得点を目指す」単目的行動モデル（高さ）、「振り込みを避ける」単目的行動モデル（安全さ）の合法手中の評価順位を用いて、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。各タイプの詳細は表 6.1 に示す。

局面の分類の際にはまず3つの単目的行動モデルによって全合法手を評価し、各打牌のそれぞれのモデルにおける評価順位を求める。（「早い和了を目指す」「高得点を目指す」「振り込みを避ける」単目的行動モデル上のある打牌 a の合法手中の評価順位をそれぞれ $\text{rank}_H(a)$, $\text{rank}_T(a)$, $\text{rank}_A(a)$ と表すこととする）

次に式 6.1 によって3つのモデルの評価順位を組み合わせた評価値 $\text{rank}(a)$ を各合法手について導出、上級者の打牌の $\text{rank}(a)$ の合法手内で上位何位に入っているかによってタイプの分類を行う。分類はタイプ 1~7 まで順に分類条件を満たすかどうかを判定し、条件を満たした場合にそのタイプであると分類し、条件を満たさなければ次のタイプへと進む、いずれのタイプにも分類されなかったものは今回使用した単目的行動モデルでは説明のできない不明タイプとし決定木の学習の際には取り除く。各タイプに分類される局面の割合の調整の結果、分類条件は、T1 が「上級者の打牌の $\text{rank}(a)$ が合法手内で単独で最小」、T2~T4 が「上級者の打牌の $\text{rank}(a)$ が合法手内で最小（最小が複数あっても分類）」、T5~T7 が「上級者の打牌の $\text{rank}(a)$ が合法手内で上位3位以内」とした。

表 6.2 に決定木の学習に使用する上級者の牌譜 122 半荘分、61771 局面の分類結果を示す。不明とされた局面が約 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。

$$\text{rank}(a) = w_H \times \text{rank}_H(a) + w_T \times \text{rank}_T(a) + w_A \times \text{rank}_A(a) \quad (6.1)$$

表 6.1: 学習データの分類法

タイプ	概要	重み (早さ (w_H), 高さ (w_T), 安全さ (w_A))
T1	全戦略において良い打牌	(1, 1, 1)
T2	早さ, 高さの観点で良い打牌	(3, 3, 1)
T3	早さ, 安全さの観点で良い打牌	(3, 1, 3)
T4	高さ, 安全さの観点で良い打牌	(1, 3, 3)
T5	早さの観点で良い打牌	(5, 1, 1)
T6	高さの観点で良い打牌	(1, 5, 1)
T7	安全さの観点で良い打牌	(1, 1, 5)
不明	説明できない打牌	—

表 6.2: 学習データのタイプ分類結果

	Total	T1	T2	T3	T4	T5	T6	T7	不明
局面数	61771	13857	9643	1581	1623	12685	8310	1480	12592
割合 (%)	100	22.4	15.6	2.6	2.6	20.5	13.5	2.4	20.4

6.2 学習方法

表 6.3 が本研究で決定木の学習に使用した特徴量の一覧である。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。

表 6.3: 決定木の学習に使用した特徴量一覧

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}
現在の順位	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}

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。

表 6.2 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。各局面の分類結果を表 6.4, 入力された局面のタイプごとの正答率を表 6.5 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切れない値である。学習データのタイプの分類時には、タイプ 1~7 と不明タイプの中で最も近いタイプに分類するのではなく、タイプ 1~7 まで順に分類されるかどうかの判別を行い、分類された時点で以降のタイプへの分類判別は行わずにタイプを決定している。そのため、分類順の遅い「ある戦略のみを重視」しているとされるタイプ 5~7 が、その戦略含む複数の戦略を重視していると分類される場合 (例えば早さのみを重視している T5 が全戦略を均等に重視の T1, 早さと高さを重視する T2, 早さと安全さを重視する T3 に分類される場合) は、タイプは一致していなくとも許容範囲内であるといえるが、今回の結果では、T6 と T7 が戦略的に関係のない T5 に分類される局面も多く、改善の余地がある。

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。

表 6.4: 決定木によるタイプ分類結果

入力\予測結果	T1	T2	T3	T4	T5	T6	T7
T1	6594	1988	267	204	3144	1431	229
T2	1892	4037	45	105	2136	1382	46
T3	975	69	67	28	315	84	43
T4	694	234	46	51	354	222	22
T5	4090	2771	179	159	3646	1703	137
T6	2175	2200	55	109	2090	1640	41
T7	837	70	67	24	341	69	72

表 6.5: 入力タイプ別の正答率

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

第7章 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、前章で示した式6.1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7.1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7.2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。

表 7.1: タイプの正答率と予測精度

	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 7.2: タイプ別上級者の打牌予測精度

元のタイプ	決定木による予想	上級者の手の予想順位			
		1 位率 (%)	2 位率 (%)	3 位率 (%)	3 位外率 (%)
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

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。

第8章 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。

謝辞

本研究を進めるに当たり，様々なご指導頂きました池田心准教授に深謝致します。

また，日常の議論を通じて多くの知識や示唆を頂いた池田研究室・飯田研究室の皆様
に感謝いたします。

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, Vol. 2007, 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] 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, ゲームプログラミングワークショップ2013 論文集, Vol. 2013, pp 1-7, 2013.
- [5] 中張遼太郎, 水上直紀, 浦晃, 三輪誠, 鶴岡慶雅, 近山隆, 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, Vol. 2013, 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, 福地誠, 『おしえて! 科学する麻雀』, 洋泉社, 2009.