

有向非巡回グラフで表現された1人麻雀の探索アルゴリズム

栗田 萌¹ 保木 邦仁²

概要: 麻雀において、他人の行動をほとんど無視して自分のアガリとテンパイのみを目指すよう簡略化されたゲームを1人麻雀と呼ぶ。本研究では、1人麻雀の探索木の節点のいくつかを同一視して、節点数を削減した有向非巡回グラフ (DAG) の探索アルゴリズムを提案する。このアルゴリズムが1人麻雀のゲーム木を探索するときに列挙する手牌の数をシャンテン数に応じて計測し、シャンテン数が4程度の場合に手牌の数は10万程度となる結果を得た。

Search Algorithm of Single-Player Mahjong Represented by Directed Acyclic Graph

MOYURU KURITA¹ KUNIHITO HOKI²

Abstract: A simplified mahjong game in which one player aims for *agari* or *tenpai* by neglecting most of the others' actions is called single-player mahjong. In this research, we propose a search algorithm of directed acyclic graph (DAG) of single-player mahjong, where DAG equates several nodes of the game tree with a node. We have measured the *shanten* number dependency on the number of the player's hands enumerated for search, and found the number of hands is about 100 thousand when the *shanten* number is about 4.

1. はじめに

麻雀は多人数不完全情報ゲームであり、1回のゲームが終わった際に各プレイヤーは点数と順位に応じた利得を得る。通常1回のゲームは複数の局からなり、1局はプレイヤーの誰かがアガリをするか、牌山が無くなるなどの理由で流局するかで終了する。1局の終わり方に応じて、プレイヤー間の点数状況が変化し、最終局終了時点で点数が多い順に1回のゲームの順位が決定する。1つの局でアガリを行うことができるのは基本的に1人のプレイヤーに限られることから、他プレイヤー（他家）よりも早くアガリの手牌を作る技術は麻雀の実力を決める重要な要素である。また、点数の高いアガリ形ほど一般的には作りやすく、プレイヤーは現在の点数の状況や、他家の手牌の進行を考慮に入れたうえで、どのようなアガリを目指すか考える必要がある。確定完全情報ゲームの場合と異なり、プレイヤーは他プレイヤーがアガリをすることができるかどうか知ることができない

ため、打牌やフーロの選択を行う際に、その選択を行った後のアガリ確率や利得期待値を高い精度で推定することは麻雀のAIを作る上で重要な課題である。

他家の行動をほとんど無視して自分のアガリだけを考える1人麻雀では、終端節点の利得を設定することで、その期待値を最大化する定式化が比較的容易になる。また、終端節点の利得を局終了時の順位点期待値とすることで、点数状況に応じたアガリを目指すことが可能になる。その一方で、1人麻雀はゲーム木のサイズが非常に大きいという難点を持つ。著者らは先行研究において、この問題を解決するため、ゲーム木を有向非巡回グラフ (DAG) を用いて近似的に表現するモデルを提案した [1]。このモデルでは、ゲーム木の1つ以上の節点を、手牌 q と打牌回数 t とゲーム進行のフェーズにもとづいて1つの DAG の節点にまとめる。ゲーム木の簡略化は麻雀のルールに基づいて行い、終端節点の利得の設定も点数のみを特徴量とした機械学習を用いて行うことができるため、麻雀の高い専門知識を直接コーディングすることなしに強い人工知能 (AI) を開発することが可能になる。先行研究では、探索の具体的なアル

¹ HEROZ 株式会社 mkmjai1@gmail.com

² 電気通信大学 k.hoki@uec.ac.jp

ゴリズムと、どの程度のサイズの DAG をどの程度の時間で探索できるかについて明らかにされていなかったため、これらを明らかにすることを本研究の目的とする。

2. 麻雀のルールと用語

この章では、麻雀のルールと用語について解説する。麻雀では 136 枚の牌を利用し、この中には 34 の牌種が 4 枚ずつ入っている。牌種には 7 つの字牌と 27 の数牌があり、数牌は 3 つの色に分類される。初めに各プレイヤーに 13 枚の牌が配られ、配られなかった牌は牌山に伏せられる。そこからルールで定められた方法で牌を 1 枚ずつ交換し、役と呼ばれる特定の構成を作り和了 (アガリ) し、役に応じた点数を得るゲームである。アガリの点数は直前の手牌と最後に入手した牌とその入手方法に依存し、牌山から 1 枚牌を引いた場合をツモアガリ、他プレイヤー (他家) が捨てた牌を利用する場合をロンアガリという。牌を交換する方法としては、牌山から 1 枚牌を引く (ツモ) 方法と、他プレイヤー (他家) が捨てた牌を入手 (フーロ) する方法がある。交換の場合は牌を入手した後不要な牌を 1 枚場に捨てる。特殊な役を除き、アガリ手牌は 1 つの対子 (トイツ) と 4 つの面子 (メンツ) とからなる。対子とは同じ牌種の 2 枚組のことであり、面子とは特定の条件を満たす 3 枚または 4 枚の牌の組である。同じ牌種 3 枚からなる面子を刻子 (コーツ) といい、連続する数字の数牌 3 枚からなる面子を順子 (シュンツ) という。これらのメンツの最後の 1 枚を他家の捨て牌からフーロにより入手することができ、このフーロを刻子の場合ポン、順子の場合チーという。フーロによって完成したメンツは他家から見えるように場に晒す。フーロによって完成した刻子を明刻 (ミンコー)、フーロせずに完成した刻子を暗刻 (アンコー) という。また、あまり一般的な語ではないがフーロによって完成した順子を明順 (ミンシュン) という。同じ牌種 4 枚からなる面子を槓子 (カンツ) という。明刻と同じ牌種を自分の手牌から場に出して槓子にする操作を加カン、暗刻から他家の捨て牌を入手して槓子とするフーロを大明カン、フーロせずに揃えた手牌の中の同一牌種 4 枚を槓子として扱うことを宣言することを暗カンという。暗カンは他家の捨牌を入手する方法ではないが、本研究ではこれもフーロと見做す。加カン、大明カンによって完成した槓子を明槓といい、暗カンにより完成した槓子を暗槓という (この論文では、フーロとしてのカンをカタカナで、面子としての槓を漢字で表記する)。

また、特定の条件下においてプレイヤーはリーチを行うことができ、これは 1000 点を支払う必要があるが役になる。この 1000 点はその局でアガリした人が得る。牌山が無くなるまで誰もアガリしなかった場合を流局といい、各プレイヤーの最後の手牌の形に応じた点数を各プレイヤーが得る。リーチしたプレイヤーがいる場合、支払われた 1000 点は次

の局でアガリしたプレイヤーの点数となる。

3. 先行研究

1 人麻雀において、現在までに提案されたアガリを目指す手法には、ニューラルネットワークやパーセプトロンなどの機械学習を用いて上級プレイヤーの行動を予測するものや [2], [3]、モンテカルロ法やタブーサーチを用いてゲームの試行や木探索をするものがある [4], [5]。機械学習を用いる手法では、ゲームの試行や木探索を行わないため意思決定に時間がかからないという長所があるものの、点数状況に応じたアガリを目指すことが難しいことが報告されている。一方で、モンテカルロ法やタブーサーチを用いて探索を行うような手法では 1 人麻雀での最適な行動を求め得るが、現在までに降りを目指す方策も考慮した実践的な AI の作成には至っていない。これは、ゲーム木のサイズが大きすぎるためにモンテカルロ法の収束が遅いことなどが原因と考えられる。

これらの課題を解決するため、著者らが先行研究で抽象化した 1 人麻雀は以下のルールを持つ。はじめにプレイヤーに 13 枚の手牌 $q \in Q$ が与えられる。プレイヤーは定められた打牌回数 t_{\max} の中で、ツモかフーロを行いながらアガリやテンパイを目指す。ゲームの進行を以下の 7 つのフェーズに分ける。

- **start:** プレイヤーに 13 枚の手牌 $q \in Q$ が与えられ、 chance_F へ移行する。
- **chance_F:** プレイヤーの打牌回数が t_{\max} 未満の場合、他家が牌を牌山から 1 枚場に出し player_F へ移行する。打牌回数が t_{\max} の場合、手牌に応じた利得が与えられてゲームが終了する。
- **player_F:** プレイヤーは手牌と他家の捨て牌に応じてフーロやロンが可能で、ロンした場合 agari_R へ移行する。大明カンした場合は chance_F へ移行し、打牌回数が 1 増えたものとして扱う (嶺上牌のツモは行わない)。大明カン以外のフーロをした場合 player_D へ移行し、フーロしない場合 chance_T へ移行する。
- **chance_T:** プレイヤーは牌山から牌を 1 枚ツモし player_D へ移行する。
- **player_D:** プレイヤーは可能なアガリまたは打牌の一つを選択し、アガリを宣言した場合には agari_T へ移行し、アガリをしない場合プレイヤーは牌を 1 枚捨てて chance_F に戻る。暗カン、加カンした場合も打牌回数が 1 増えたものとして chance_F に戻る。
- **agari_T, agari_R:** ルールに基づいた利得が与えられてゲームが終了する。

これらのゲームの進行は図 1 のようにまとめられる。

このゲームの部分木は DAG として簡略化することができて、利得の期待値は手牌 q 、ツモ、フーロ可能牌種 h 、フェーズ a 、打牌回数 t 、根節点 n_0 の関数として近似的に

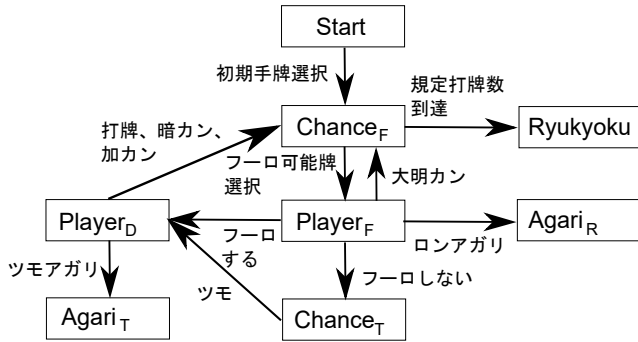


図 1 1人麻雀のゲーム進行

表される。なお、1人麻雀の木の根節点に対応する DAG の節点も、根節点と呼ぶことにする。アガリをした場合、すなわち $a \in \{\text{agari}_T, \text{agari}_R\}$ ならば

$$E(q, h, a, t, n_0) = U_{\text{agari}}(q, h, a, n_0) \quad (1)$$

となる。ここで、 $U_{\text{agari}}(q, h, a, n_0)$ は現在の点数状況、アガリを行う直前の手牌 q 、アガリ牌 h 、アガリがツモかロンかを表すフェーズ a などに依存する利得である。

ツモまたはフーロの後のプレイヤー節点、すなわち $a = \text{player}_D$ が成り立つとき

$$E(q, i, \text{player}_D, t, n_0) = \max_{(q_c, i_c, a_c) \in C(q, i, \text{player}_D)} E(q_c, i_c, a_c, t + 1, n_0) \quad (2)$$

となる。ここで、 i はこの節点の直前のツモまたはフーロを表す牌種である。また、 $C(q, i, \text{player}_D)$ は、手牌 q においてツモまたはフーロ i を行った後に行うことができる打牌またはアガリ宣言を行った後の全ての手牌 q_c 、ツモまたはフーロ i_c 、フェイズ a_c の組からなる集合である。

フーロ、ロンの選択を行う節点、すなわち $a = \text{player}_F$ であるとき

$$E(q, h, \text{player}_F, t, n_0) = \max_{(q_c, h_c, a_c) \in C(q, h, \text{player}_F)} E(q_c, h_c, a_c, t, n_0) \quad (3)$$

となる。ここで、 $C(q, h, \text{player}_F)$ は、手牌 q において他家が牌種 h を打牌した後に行うことができるフーロ、ロンを行った後、またはフーロもロンも行わなかった全ての場合における手牌 q_c 、フーロ h_c 、フェイズ a_c の組からなる集合である。

また偶然節点については、ツモ前の節点と他家打牌前の節点が存在し、それぞれ

$$\begin{aligned} & E(q, \text{null}, \text{chance}_T, t, n_0) \\ &= \sum_{h \in H} p_T(q, h, t, n_0) E(q, h, \text{player}_D, t, n_0) \\ & \quad E(q, \text{null}, \text{chance}_F, t, n_0) \\ &= \sum_{h \in H} p_F(q, h, t, n_0) E(q, h, \text{player}_F, t, n_0) \end{aligned} \quad (4)$$

と書ける。ここで、 $p_T(q, h, t, n_0)$ と $p_F(q, h, t, n_0)$ はそれぞれ、牌種 h をツモする確率と他家が捨ててフーロできる確率である。

4. 提案手法

本章では、DAG を用いて近似的に表現した 1 人麻雀のデータ構造と探索アルゴリズムを提案する。目標は、一般的な対戦条件（例えば数秒程度の思考時間）のもとで局をプレイ可能な麻雀 AI に組み込むことである。

4.1 手牌集合の表現

1 局は、牌のツモ及びフーロと打牌の繰り返しで、手牌が合計 13 枚の牌からなる節点と、14 枚の牌からなる節点が存在する。合計 13 枚の牌からなる手牌全ての集合を Q 、合計 14 枚の牌からなる手牌全ての集合を Q' とする。ここでの合計枚数は、フーロによって場に晒された牌も含めるものとする。例外的にカンツは 4 枚の牌から構成されるが、便宜的にこれを 3 枚と数えて手牌集合を Q および Q' に分割する。はじめに手牌 $q \in Q$ と $q' \in Q'$ が持つ情報を述べる。手牌は以下の要素で構成される。

- 非公開牌種：その手牌を持つプレイヤーのみに見える場に晒していない情報であり、34 種の牌をそれぞれ何枚持っているかを表す整数の組。各要素は 0 から 4 までの整数値を取り、要素の総和はフーロの回数を m_f として $13 - 3m_f$ または $14 - 3m_f$ である。
- フーロ：明順 21 種、明刻 34 種、暗槓 34 種、明槓 34 種が存在し、それぞれの個数を表す整数の組。明順に関する 21 次元の要素は 0 から 4 までの整数を取り、それ以外の要素は 0 または 1 の整数により表される。
- リーチ：リーチを表すフラグ。
- 赤ドラ：赤ドラは 3 枚のみであり、それぞれについて、持っていない場合、持っていて場に晒していない場合、持っていて場に晒している場合の 3 通りの状態を持つ。

手牌に関する上記に含まれない情報については省略され、それらが異なる手牌は同一のものとして扱う。具体的には、フーロを行った順番と入手元他家の情報や、チーにおいてフーロした牌種の情報、大明カンと加カンの違い、赤ドラを含むフーロの種類情報が省略されている。

実装においては上に挙げた情報はいくつかの 32 ビット整数型により表される。具体的には、非公開牌種は

$$\text{uint32}_t \text{ color}[4] \quad (5)$$

を用いて表現する。各牌種は 4 枚までしか持てないため、1 つの牌の枚数は 3 ビットで表現できる。従って、一つの色（萬子、索子、筒子）の各牌の枚数は 27 ビットのデータ構造で、字牌については 21 ビットのデータ構造で表現可能である。このように定義すると、ある手牌からの牌種の増減を整数の足し算と引き算で表現することが可能で

ある。実際には、合計枚数に制約があるため、より少ないビット数で非公開牌種を表現可能だが、簡単のためこのようにした。以下では、この 32 ビット整数型 4 つをまとめて Hai_Num_Bit と呼ぶ。

後に探索の対象とする手牌を列挙する段階において、同じ手牌を複数回探索しないために Hai_Num_Bit に対してハッシュ値を定義すると便利である。本研究では XOR と定数シフトのみの演算からなる Algorithm 1 を用いた。記号 \ll は、ビット列の左シフトを表している。多くの場合麻雀の手牌は 1 種類の牌を 0 または 1 枚持つことが多いため、色に合わせて左シフトの数を変更することで、ハッシュが衝突する確率を減らすことが可能になる。

Algorithm 1 非公開牌種のハッシュ関数

```

result ← 0
for i ∈ [0, 3] do
  result ← result XOR (color[i] ≪ i)
end for
return result

```

フーロ、リーチ、赤ドラの情報もいくつかの 32 ビット整数を用いて表現することができる。実際にこれらは

```

uint32_t minshun_bit[3],
uint32_t minko_kan_bit[4],
uint32_t reach_aka_bit

```

(6)

の整数 8 つで表現可能である。明順については、一つの整数が一つの色に対応していて、一つの色に 7 種類のチーが存在して、同種のもは 4 つまで存在し得るため、 $7 \times 3 = 21$ ビットで表すことができる。明刻、明槓、暗槓については、同じ牌のものは 1 つまでしか存在し得ず、各色に明刻、明槓、暗槓合わせて 27 種存在するため、27 ビットで表すことができる。字牌については合計 21 種存在するため、21 ビットで表現できる。最後にリーチと赤ドラについては、リーチは 1 ビットで表し、各色の赤が晒されているかどうかを 1 ビット（合計 3 ビット）で表し、手牌の中に持っているかどうかも 1 ビットで表した。これらについてもハッシュ値を定義すると都合がよく、本研究では 8 つの整数の XOR 積をハッシュ値とした。今後、これら 8 つの 32 ビット整数型をまとめて、Tehai_State_Bit と呼ぶ。

4.2 牌交換の表現

続いて、ある手牌から任意の回数だけ牌をツモし、打牌を行う操作について考える。本研究ではこの動作を牌交換と呼び、牌交換全てからなる集合を C により表す。牌交換は、手牌の各牌種の増減数を要素とする 34 個の整数によって表すことができ、各要素は -4 から 4 までの整数値をとる。 C にはルール上許されないような交換（例えば全要素 -4 により表現される交換）も属す。

牌交換 $c \in C$ に対して、牌種 $h \in H$ の増減を c_h で表す。また、手牌に増えた牌の数を n 、減った牌の数を m として

$$C(n, m) = \left\{ c \in C : \begin{aligned} n &= \sum_{h \in H} \max(c_h, 0), \\ m &= \sum_{h \in H} \max(-c_h, 0) \end{aligned} \right\} \quad (7)$$

とする。プレイヤーが打牌選択を行う際、手牌は合計 14 枚であり、そこから 13 枚と 14 枚の手牌を繰り返す、アガリ形となる 14 枚の手牌を目指すため、打牌選択において重要な集合は $C(n, n)$ と $C(n, n+1)$ である。以下では手牌 q において、牌交換 c を行うことで実現する手牌を $q+c$ と書く。

4.3 DAG の節点の列挙

ある打牌選択を行う節点を DAG の根節点として、探索対象となる手牌の集合を定義する。この DAG の各節点は、手牌 q 、牌種 h 、フェーズ a 、打牌回数 t により特定される。これらの内、手牌 q 以外の値は列挙することが容易である。従って、本節では手牌の列挙についてのみ述べる。4.3.1 項ではツモと打牌のみを行なった場合での探索手牌集合 Q_{S_0} を導入する。続いて、4.3.2 項では、ツモに加えてフーロにより牌を入手したり、リーチやカンを行ったりもする場合での探索手牌集合 Q_S を導入する。

4.3.1 探索手牌（ツモのみ）の列挙

1 人麻雀の DAG の探索は、想定されるテンパイ形を列挙することからはじまる。この打牌選択を行う節点での手牌を q'_0 とし、手牌をアガリ形に変える牌交換からなる集合 C_A と、 n 枚以下でアガリ形に変える牌交換からなる集合 $C_A(n)$ を以下で定義する。

$$C_A = \{c \in C : q'_0 + c \text{ がアガリ形}\}$$

$$C_A(n) = \left(\sum_{n'=1}^n C(n', n') \right) \cap C_A \quad (8)$$

$C_A(n)$ に属する要素の列挙は、シャンテン数を数えるバックトラック法を利用して、シャンテン数が $n-1$ 以下になる面子切り分け方法を列挙することによりなされる*1。

次に、探索を行う手牌を列挙するため C の要素に対して部分交換を定義する。牌交換 c' が c の部分交換であるとは以下の条件を満たすものをいう。

- $0 \leq c_h$ なる $h \in H$ に対して $0 \leq c'_h \leq c_h$ が成り立つ。
- $0 \geq c_h$ なる $h \in H$ に対して $0 \geq c'_h \geq c_h$ が成り立つ。

ここで、 c 自身も c の部分交換であり、 c の部分交換全てからなる集合を $C_P(c)$ と定義する。アガリを目指す際に探索すべき手牌を生成する牌交換の集合 $C_S(n)$ は以下のように定義される。

*1 麻雀 C 言語プログラム集。
<http://cmj3.web.fc2.com/>

$$C_S(n) = \left(\bigcup_{c \in C_A(n+1)} C_P(c) \right) \cap \left(\sum_{n'=0}^n C(n', n'+1) \right) \quad (9)$$

また、打牌選択を行う節点での全合法打牌の集合を C_L とする。これは $C(0,1)$ の部分集合であり、要素数は 14 以下である。これにより、探索の対象となる手牌全てからなる集合 $Q_{S0}(n)$ を

$$Q_{S0}(n) = \{q'_0 + c : c \in C_S(n) \cup C_L\} \quad (10)$$

と定義する。交換枚数 n は q'_0 のシャンテン数以上の値を取る必要があり、 n とシャンテン数の差を大きくとるほど探索の精度は向上する。以下では n は定数を想定するため、 $C_S(n), Q_{S0}(n)$ は C_S, Q_{S0} と略記する。

C_S の要素の中には複数の $C_P(c)$ の要素となっているものが数多く存在する。実装において重複無く Q_{S0} の列挙を行うにあたり、 Q_{S0} の要素に対する Hai.Num.Bit のハッシュ値を用いると効率が良い。

4.3.2 探索手牌（フーロ等も考慮）の列挙

今まで、牌の交換などは非公開牌種の枚数の増減のみを考えてきた。そのため、ここまでで想定できる手牌は現在の手牌 q'_0 からツモによる打牌交換を行った手牌だけであり、フーロにより実現する手牌は Q_{S0} に属さない。そこで、フーロを含めて牌の増減が c で表される手牌について考える。

牌交換 c は、手牌に増える牌と減る牌の双方を 34 個の整数組で表現できる。これらを、それぞれ、 $\mathbf{x}_{in}(c), \mathbf{x}_{out}(c)$ と書く。これらの各要素は 0 から 4 の整数である。そこでポンとチーも 34 個の整数組で表すことを考える。ポンとチーにより牌を得る場合、チーは同じ牌種のフーロでも順子の数の小さい側に入れるか、中間に入れるか、大きい側に入れるかによって異なる手牌になるため、ポンとチーによる牌の増加を 4 つの 34 個の整数組 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ により表す。ここで、 $\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ はチーする牌種を表して、字牌の部分は常に 0 であるが、ポンと合わせるため 34 個の整数組とする。結果として、 \mathbf{x}_1 の要素の総和は根節点からポンした回数、 $\mathbf{x}_2 + \mathbf{x}_3 + \mathbf{x}_4$ の要素の総和はチーした回数に対応する。このとき、フーロを含めて牌の増減が c で表される交換の集合 $D(c)$ を

$$D(c) = \left\{ (\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_{out}(c)) \right. \\ \left. : \forall h \in H, \sum_{i=0}^4 x_{ih} = x_{inh}(c), x_{ih} \in \{0, 1, 2, 3, 4\} \right\} \quad (11)$$

により定義する。さらに、 $D(c)$ の部分集合でフーロの合計回数が m_{f-max} 以下からなるフーロも考慮した交換からなる集合を $D(c, m_{f-max})$ と書く。

手牌 q'_0 から $D(c, m_{f-max})$ の交換を行った手牌からな

る Q の部分集合を $Q_D(q, m_{f-max})$ とする。すなわち、 $q = q'_0 + c \in Q_{S0}$ に対して、

$$Q_D(q, m_{f-max}) = \{q'_0 + d : d \in D(c, m_{f-max}), q'_0 + c = q\} \quad (12)$$

である。ここで $q'_0 + d$ とは、手牌 q'_0 からフーロも考慮した交換 d を行って実現される手牌である。ツモの場合と異なり、フーロの場合入手できる牌には麻雀のルールに基づいた制限があり、 $Q_D(q, m_{f-max})$ から実現不可能なものがある程度除外する。制限とは例えば、「ポンする際には同じ牌が手牌に 2 枚必要」などである。本研究では、 q'_0 と \mathbf{x}_0 から判定可能なルール上の制約を考える。この条件は一般に以下のようにまとめられる。フーロ $\mathbf{x}_i, i \in \{1, 2, 3, 4\}$ を行うために必要な牌種と枚数を表す整数の組を \mathbf{y}_i とする。例えば \mathbf{x}_1 が牌種 h のポン 1 つからなる場合 $y_{1h} = 2$ で他は 0 である。 Q_D から除外される手牌は、ある牌種 h に対して

$$q'_{0h} + x_{0h} < \sum_{i=1}^4 y_{ih} \quad (13)$$

が成り立つものである。ここで q'_{0h} は手牌 q'_0 の非公開牌種に含まれる牌種 h の枚数である。

さらに、探索する手牌に、リーチをした手牌や、カンを含む手牌を加えることを考える。そのため、集合 $Q_D(q, m_{f-max})$ に含まれる各手牌に対して以下の手牌も考える。

- フーロ回数が m_{f-max} を超えない範囲で、手牌に含まれる暗刻を暗槓または明槓に置き換えたり、明刻を明槓に置き換えたりした手牌（フーロ回数は場に晒した面子の数であり、加カンはフーロ回数を増やさないものとする）。
- $Q_D(q, m_{f-max})$ の要素と上で挙げたメンゼンテンパイ手牌においてリーチをした手牌。

これらの手牌を $Q_D(q, m_{f-max})$ に加えた集合を $Q'_D(q, m_{f-max})$ とする。

Q_{S0} の全ての要素に対する $Q'_D(q, m_{f-max})$ の和集合が探索する手牌の全体であり、 Q_S と書く。すなわち

$$Q_S = \sum_{q \in Q_{S0}} Q'_D(q, m_{f-max}) \quad (14)$$

である。以降では、 $Q'_D(q, m_{f-max})$ を $Q'_D(q)$ と略記し、その要素を q の派生手牌と呼ぶ。

実装においては、上述した手牌の交換を再帰的に行い、今までに無い手牌が現れたらそれを集合に追加する。これらは $q_i \neq q_j$ なる $q_i, q_j \in Q_{S0}$ に対して $Q'_D(q_i) \cap Q'_D(q_j) = \emptyset$ が成り立つため、各 Q_{S0} の要素間で並列に実行可能である。また、 $Q'_D(q)$ の各要素は Tehai.State.Bit により識別可能であるから、このハッシュ値を利用することで重複のない $Q'_D(q)$ の要素列挙が効率的に行われる。

4.4 枝の列挙

前節までで、 q'_0 からアガリに向かう際に、牌交換回数とフーロ回数を閾値として探索空間に制限を課し、手牌の列挙を行った。実際に DAG の探索を行う際には、枝の情報も必要である。はじめに、フーロを行わない各手牌に対して、牌交換を行った手牌が Q_{S0} の要素であるような一枚牌交換 $C(1, 1)$ の部分集合を定義する。すなわち、 $q \in Q_{S0}$ に対して $e(q)$ を

$$e(q) = \{c \in C(1, 1) : q + c \in Q_{S0}\} \quad (15)$$

と定義する。 $C(1, 1)$ に属する要素の数は 34×33 通りであり、 $e(q)$ の列挙の実装は、 q に対し $C(1, 1)$ の全ての要素に対し $q + c$ を計算し、それが Q_{S0} に含まれるかを判定する。これも Hai_Num_Bit のハッシュを利用すると、ハッシュ表が大きい理想的な状況では計算量を $O(1)$ に抑えることが可能である。また、この計算は各 $q \in Q_{S0}$ に対して独立に実行可能であるため並列化が容易である。

前節でフーロ等を含めて手牌を列挙したが、これに合わせて1枚の牌交換に関しても、フーロを含めて拡張する必要がある。牌種 h_{in} を1枚増やし、牌種 $h_{out} (\neq h_{in})$ を1枚減らす交換 $c \in C(1, 1)$ に対して、以下の交換からなる集合を派生交換集合 $A_1(c)$ とする。

- ツモ： h_{in} をツモし、 h_{out} を打牌する。リーチしない場合とする場合の2通りが存在する。
- チー： h_{in} をチーし、 h_{out} を打牌する。順子のどの部分をチーするかによって3通り想定される。
- ポン： h_{in} をポンし、 h_{out} を打牌する。
- 加カン、暗カン： h_{in} をツモし、 h_{out} を加カンまたは暗カンする。

手牌 $q \in Q_S$ に対して、交換 $d \in A_1(c)$ を行うことで実現する手牌を $q + d$ とする。DAG の探索を行う際、各 $q \in Q_S$ に対して、 $q + d \in Q_S$ が成り立つような $d \in A_1(c)$ 全てを予め列挙すると効率が良い

ここで $q_i \in Q_{S0}$ 、 $c \in e(q_i)$ 、 $q_j \in Q'_D(q_i)$ 、 $d \in A_1(c)$ で

$$q_j + d \in Q_S \Leftrightarrow q_j + d \in Q'_D(q_i + c) \quad (16)$$

という性質が成り立つ。したがって、 $q_j + d$ の Tehai_State_Bit のハッシュ値を利用することで効率的に DAG の探索に必要な派生交換を列挙することが可能である。

続いて牌種 h の自己交換 $A_0(h)$ を以下で定義する。

- 加カン、暗カン：牌種 h をツモし、牌種 h を加カンまたは暗カンする。
- 大明カン：牌種 h を大明カンする。

これらは集合 $Q'_D(q)$ の中で閉じた交換である。すなわち、 $q_j \in Q'_D(q_i)$ 、 $d \in A_0(h)$ であるとき

$$q_j + d \in Q_S \Leftrightarrow q_j + d \in Q'_D(q_i) \quad (17)$$

が成り立つ。ツモ切りリーチは集合 $Q'_D(q)$ の中で閉じた交換であるから、この中に含まれ得るものであるが、フーロやリーチを行う他家がない状況で、リーチを数巡待つことが有利になることは想定しづらいことから本研究では考えないものとする。

手牌 q に対して列挙した交換の集合を $A(q)$ とする。すなわち、

$$A(q) = \left\{ d \in \sum_{c \in C(1, 1)} A_1(c) \cup \sum_{h \in H} A_0(h) : q + d \in Q_S \right\} \quad (18)$$

である。

これらの列挙のアルゴリズムは Algorithm 2 にまとめられる。1行目の Q_{S0} に関するループは、そのサイズが数万程度であり他のループに比べて一桁以上大きいため並列化して行う。2行目の $Q'_D(q_{S0})$ に関するループのサイズは、探索を始める時の手牌のシャンテン数や、フーロ回数に依存するが、およそ10以下である。3行目の $e(q_{S0})$ のループのサイズは $|C(1, 1)|$ より小さく、実際は数百程度である。4行目の $A_1(c)$ については、入手する牌がチーできる牌かどうかなどによってサイズが異なるが、10以下であり、11行目の $A_0(h)$ のサイズは3である。5行目と12行目で、手牌 q から変換 d を行った手牌 $q + d$ が Q_S に属するか判定を行う。この際、 Q_S はサイズが10万程度の集合であるが、式(16)、(17)の関係を用いると、判定はおよそ10以下の集合から行えばよいことがわかる。

Algorithm 2 DAG の枝列挙アルゴリズム

```

1: for  $q_{S0} \in Q_{S0}$  do in parallel
2:   for  $q \in Q'_D(q_{S0})$  do
3:      $A(q) \leftarrow \emptyset$ 
4:     for  $c \in e(q_{S0})$  do
5:       for  $d \in A_1(c)$  do
6:         if  $q + d \in Q'_D(q_{S0} + c)$  then
7:            $A(q) \leftarrow A(q) \cup \{d\}$ 
8:         end if
9:       end for
10:    end for
11:    for  $h \in H$  do
12:      for  $d \in A_0(h)$  do
13:        if  $q + d \in Q'_D(q_{S0})$  then
14:           $A(q) \leftarrow A(q) \cup \{d\}$ 
15:        end if
16:      end for
17:    end for
18:  end for
19: end for

```

4.5 終端節点の利得

手牌の形によっては、あと1枚の牌でアガリ宣言を行うことができる。アガリ宣言を行った場合、プレイヤーにはアガリ形に応じた利得が与えられゲームは終了する。手牌

$q \in Q_S$ においてアガりに必要な牌と利得 $(h, u) \in H \times \mathbb{R}$ からなる集合を、ツモアガリとロンアガリのそれぞれで、 $B_{\text{tsumo}}(q), B_{\text{ron}}(q)$ とする。また、探索を行う上での最大想定打牌回数 t_{max} の打牌を行ったときに手牌が q であった場合の利得を $U_{\text{ryukyoku}}(q)$ とする。

4.6 偶然節点における行動の確率

DAG の探索を実際に行うためには、偶然節点における行動の確率を定める必要がある。これは、ある牌 h をツモする確率および副露できる確率であり、以下のようにまとめることができる。

- ツモ確率：牌山から牌種 h をツモする確率であり、打牌回数 t とその時の手牌 q に依存するとして $p_T(q, h, t)$ と表す。
- チー確率：上家が牌種 h を打牌する確率であり、打牌回数 t とその時の手牌 q に依存するとして $p_F(q, h, t)$ と表す。
- ポン、ロン確率：他家が牌 h を打牌する確率であり、チー確率の3倍と考えて $3p_F(q, h, t)$ とする。

4.7 期待値の計算

前節までで、 Q_S の各要素に対して、どの行動を行った後にどの手牌に遷移するかを列挙した。ここでは、その結果を用いて1人麻雀の期待値の計算アルゴリズムを説明する。そのため、 $A(q)$ をツモにより行うことができる行動の集合 $A_{\text{tsumo}}(q)$ 、チーの集合 $A_{\text{chi}}(q)$ 、ポンと大明カンの集合 $A_{\text{pon}}(q)$ とに分割する。また、 $E(q, \text{null}, \text{chance}_F, t, n_0), E(q, \text{null}, \text{chance}_T, t, n_0)$ はそれぞれ $E(q, F, t), E(q, T, t)$ と略記する。具体的な計算は Algorithm 3 に従う。 $E(q, F, t)$ の計算に用いる $flag(h)$ は、その牌を他家から手に入れる最適な行動がポン、大明カン、ロンである場合に1となり、フーロの確率の係数を3に変える効果がある。

探索の具体的な順序は以下のようなものである。はじめに打牌回数が t_{max} となった終端節点の利得とアガリ確率を設定する。この節点では、1人麻雀のルールによりツモもフーロもできないことから、アガリ確率は0であり、手牌に応じた利得 $U_r(q)$ が設定される。

次に、終端節点 ($t = t_{\text{max}}$) から現在の節点で打牌を行った節点 ($t = 1$) まで期待値とアガリ確率を更新するループ処理を行う。期待値とアガリ確率は終端節点に近いほうから確定していくため、 t が同じ節点ではフェーズがツモの節点が決まり、フェーズがフーロの節点が後に続く。これは1人麻雀のルールで、打牌を行った後にフーロのフェーズが先にあり、次にツモのフェーズが続くためである。具体的な更新は以下のように行う。牌種 h をツモした際に最も期待利得の高い選択を行った場合の期待利得とアガリ確率を $E_{\text{tmp}}(h), p_{\text{tmp}}(h)$ に入力する。これは、はじめにツモ

切りを行った場合の期待利得とアガリ確率 $E(q, F, t + 1), p(q, F, t + 1)$ を初期値として、手牌 q から行うことができるツモアガリの集合 $B_t(q)$ および、ツモによる牌交換の集合 $A_{\text{tsumo}}(q)$ から期待利得最大の選択を検索することで行う。その後、求めた $E_{\text{tmp}}(h), p_{\text{tmp}}(h)$ に対して、ツモ確率 $p_T(q, h, t)$ との積を足し合わせてツモフェーズの期待利得およびアガリ確率の更新処理が完了する。フーロフェーズの更新では、初期値をフーロを行わなかった場合の期待利得とアガリ確率 $E(q, T, t), p(q, T, t)$ に設定する。続いて、ツモフェーズの更新と同様に、各牌種が打牌された際に最も期待利得の高い選択を行った場合の期待利得とアガリ確率を $E_{\text{tmp}}(h), p_{\text{tmp}}(h)$ に入力する。これは、手牌 q から行うことができる行動の集合 $A_{\text{chi}}(q), B_{\text{ron}}(q), A_{\text{pon}}(q)$ から検索する。この時、抽象化した1人麻雀では他家3人の打牌が1つにまとめられるが、元のルールではチーは上家からしかできないものの、ポンとロンは3人の捨て牌から行うことであることを反映するため、ポンまたはロンが最適な行動である場合はそのことも記録する。更新は、各牌種 h に対して初期値に $E_{\text{tmp}}(h) - E(q, T, t)$ および、 $p_{\text{tmp}}(h) - p(q, T, t)$ とフーロ、ロンができる確率の積を足していくことで行う。その確率は、最適な行動がチーである場合は $p_F(q, h, t)$ であり、ポン、大明カン、ロンである場合は $3p_F(q, h, t)$ である。

5. 評価実験

この章では、前節で提案した手法を用いて DAG を構成した場合に DAG の大きさがどの程度になるか、実験によって明らかにする。DAG の大きさは、根節点での手牌のシャンテン数、式 (10) の交換枚数 n 、そして式 (12) のフーロ回数 $m_{t\text{-max}}$ に大きく依存する。これらの条件で分類を行った場合の、 $|Q_S|$ の平均的な値を表1に示す。手牌は天鳳の鳳凰卓においてあるプレイヤーが4回目の打牌を行う際の手牌をランダムに選択した。シャンテン数、交換枚数、フーロ回数が大きくなるほど $|Q_S|$ の値が大きくなるが、3シャンテン、4シャンテンの場合でも交換枚数が4までであればおよそ10万程度に収まる結果を得た。この程度のサイズの DAG であれば、Intel Core i7(物理コア数4、論理コア数8)、メモリ8GBの計算機を用いておよそ5秒以下で結果を出すことが可能である。3シャンテンまでであれば、交換枚数がシャンテン数より大きくても探索が可能であり、状況に応じてシャンテン数を落とす選択が可能である。1シャンテンの場合3枚変えも探索可能であり、ターツを落として高い手を狙う選択もあり得る。

6. おわりに

本研究では、1人麻雀の探索木の節点のいくつかを同一視して、節点数を削減した有向非巡回グラフ (DAG) の探索アルゴリズムを提案した。シャンテン数に応じてアガリま

Algorithm 3 期待利得とアガリ確率の計算

```

1: for  $q \in Q_S$  do in parallel
2:    $E(q, T, t_{\max}) \leftarrow U_r(q)$ ,  $E(q, F, t_{\max}) \leftarrow U_r(q)$ 
3:    $p(q, T, t_{\max}) \leftarrow 0$ ,  $p(q, F, t_{\max}) \leftarrow 0$ 
4: end for
5: for  $t : t_{\max} - 1$  to 1 do
6:   for  $q \in Q_S$  do in parallel
7:      $E(q, T, t) \leftarrow 0$ ,  $p(q, T, t) \leftarrow 0$ 
8:     for  $h \in H$  do
9:        $E_{\text{tmp}}(h) \leftarrow E(q, F, t + 1)$ ,  $p_{\text{tmp}}(h) \leftarrow p(q, F, t + 1)$ 
10:    end for
11:    for  $(h, u) \in B_t(q)$  do
12:      if  $u > E_{\text{tmp}}(h)$  then
13:         $E_{\text{tmp}}(h) \leftarrow u$ ,  $p_{\text{tmp}}(h) \leftarrow 1$ 
14:      end if
15:    end for
16:    for  $d \in A_{\text{tsumo}}(q)$  do
17:      if  $E(q + d, F, t + 1) > E_{\text{tmp}}(h_{\text{in}}(d))$  then
18:         $E_{\text{tmp}}(h_{\text{in}}(d)) \leftarrow E(q + d, F, t + 1)$ 
19:         $p_{\text{tmp}}(h_{\text{in}}(d)) \leftarrow p(q + d, F, t + 1)$ 
20:      end if
21:    end for
22:    for  $h \in H$  do
23:       $E(q, T, t) \leftarrow E(q, T, t) + p_T(q, h, t)E_{\text{tmp}}(h)$ 
24:       $p(q, T, t) \leftarrow p(q, T, t) + p_T(q, h, t)p_{\text{tmp}}(h)$ 
25:    end for
26:  end for
27:  for  $q \in Q_S$  do in parallel
28:     $E(q, F, t) \leftarrow E(q, T, t)$ ,  $p(q, F, t) \leftarrow p(q, T, t)$ 
29:    for  $h \in H$  do
30:       $E_{\text{tmp}}(h) \leftarrow E(q, T, t)$ ,  $p_{\text{tmp}}(h) \leftarrow p(q, T, t)$ 
31:       $flag(h) \leftarrow 0$ 
32:    end for
33:    for  $d \in A_{\text{chi}}(q)$  do
34:      if  $E(q + d, F, t + 1) > E_{\text{tmp}}(h_{\text{in}}(d))$  then
35:         $E_{\text{tmp}}(h_{\text{in}}(d)) \leftarrow E(q + d, F, t + 1)$ 
36:         $p_{\text{tmp}}(h_{\text{in}}(d)) \leftarrow p(q + d, F, t + 1)$ 
37:      end if
38:    end for
39:    for  $(h, u) \in B_r(q)$  do
40:      if  $u > E_{\text{tmp}}(h)$  then
41:         $E_{\text{tmp}}(h) \leftarrow u$ ,  $p_{\text{tmp}}(h) \leftarrow 1$ ,  $flag(h) \leftarrow 1$ 
42:      end if
43:    end for
44:    for  $d \in A_{\text{pon}}(q)$  do
45:      if  $E(q + d, F, t + 1) > E_{\text{tmp}}(h_{\text{in}}(d))$  then
46:         $E_{\text{tmp}}(h_{\text{in}}(d)) \leftarrow E(q + d, F, t + 1)$ 
47:         $p_{\text{tmp}}(h_{\text{in}}(d)) \leftarrow p(q + d, F, t + 1)$ 
48:         $flag(h_{\text{in}}(d)) \leftarrow 1$ 
49:      end if
50:    end for
51:    for  $h \in H$  do
52:       $E(q, F, t) \leftarrow E(q, F, t)$ 
53:       $+ (1 + 2flag(h))p_F(q, h, t)(E_{\text{tmp}}(h) - E(q, T, t))$ 
54:       $p(q, F, t) \leftarrow p(q, F, t)$ 
55:       $+ (1 + 2flag(h))p_F(q, h, t)(p_{\text{tmp}}(h) - p(q, T, t))$ 
56:    end for
57:  end for
58: end for

```

表 1 1人麻雀を表現した DAG を構成する手牌集合のサイズの、シャンテン数、交換枚数、フーロ枚数依存性。

シャンテン数	交換枚数 n	フーロ回数 $m_{f-\max}$	$ Q_S $ 平均値
1	1	0	57
1	1	1	180
1	2	0	2323
1	2	1	5804
1	3	0	13631
1	3	1	32859
2	2	0	337
2	2	1	765
2	2	2	949
2	3	0	7865
2	3	1	20044
2	3	2	26602
3	3	0	2091
3	3	1	4802
3	3	2	6305
3	4	0	36919
3	4	1	103249
3	4	2	139468
4	4	0	10737
4	4	1	30243
4	4	2	40614

での探索をする際に列挙する手牌の数を計算し、3、4 シャンテンまでで節点数が現実的になる結果を得た。3 シャンテンまでは、必要に応じてシャンテン数を落として高い手を狙う選択が可能である。

強い麻雀の AI を構築するためには、他家の動向を無視してアガリとテンパイを目指すだけでは不十分である。特に、他家への放銃を避けることを目的とした降りの方策、アガリとテンパイと放銃の各々の確率と利得から合法手の選択を行う包括的な方策は重要である。また、シャンテン数が大きい場合の打牌方策も必要である。著者らが先行研究において提案した放銃確率の推定結果を用いた、これらの方策を含めた AI の構築は今後の課題とする。

参考文献

- [1] 栗田萌, 保木邦仁. 1人麻雀の有向非巡回グラフを用いた近似表現. 情報処理学会研究報告. GI, Vol. 2017-GI-37, No. 14, pp. 1-8, 2017.
- [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第 12 回ゲームプログラミングワークショップ, pp. 76-83, 2007.
- [3] 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四人麻雀プログラムの実現情報処理学会論文誌 Vol.55 No.11 pp.111, 2014.
- [4] 小松智希, 成澤和志, 篠原歩. 役を構成するゲームに対する効率的な行動決定アルゴリズムの提案情報処理学会研究報告. GI, Vol. 2012-GI-28, No. 8, pp. 1-8, 2012
- [5] 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, pp. 73-80, 2016.