

Transformerを用いた麻雀における手牌推定

大神卓也^{1,a)} 奈良亮耶¹ 天野克敏² 今宿祐希¹ 鶴岡慶雅^{3,b)}

概要: 麻雀の上級者は、自分から見える情報である、相手の捨て牌や自分の手牌をもとに、自分からは見えない情報である相手の手牌や牌山をある程度推測することができる。しかし、これまで研究されてきた麻雀エージェントはこれらの不完全情報の推定を明示的に行ってはいない。また、これらの推定を精度良く行うために必要な相手の行動の履歴について、その順序の情報を保持することができるアーキテクチャを用いていない。本研究では、麻雀の盤面情報を系列データとして表現し、系列データを処理することができるモデルである Transformer を用いて相手の手牌を自己回帰的に推定した。その結果、プレイヤーから見えている牌を考慮することでランダムに手牌を推定する場合に比べて高い精度を出すことができた。

Opponent Hand Estimation in Mahjong using Transformer

TAKUYA OGAMI^{1,a)} RYOYA NARA¹ KATSUTOSHI AMANO² YUKI IMAJUKU¹ YOSHIMASA TSURUOKA^{3,b)}

Abstract: Expert mahjong players use the information they can see to make some inferences about the information they cannot see, such as the opponent's hand or deck of tiles. However, the mahjong agents studied so far do not explicitly estimate such incomplete information. Moreover, they do not use architectures that can maintain information about the history of the opponent's actions, which is necessary for an accurate estimation of such information. In this study, we use Transformer, a model that can represent and process mahjong board information and estimate the opponent's hand in an autoregressive manner. As a result, we were able to achieve higher accuracy than the method that estimates the hand by randomly selecting tiles that are not visible to the player.

1. 導入

不完全情報ゲームである麻雀には、プレイヤーから見えない情報として、牌山と他プレイヤーの手牌がある。麻雀において良い選択を行うためには、プレイヤーから見える情報を利用して見えない情報の一部を推測し、選択に反映していくことが必要不可欠である。

これまで多くの麻雀エージェント [1], [2] が作成されており、中でも Suphx [1] は人間の上級者を上回る成績を残している。しかしこれらのエージェントは不完全情報の推定を明示的に行っていない。本研究では、麻雀における不完全情報の推定、特に相手の手牌を推定することを目指す。相手の手牌は役、聴牌しているかどうかなど、その他の不完全情報を包含しているため、手牌の推定を精度良く行うことができれば、その他の不完全情報の推定に活用することができる。また、相手の手牌を推定することができれば、麻雀エージェントの思考過程の可視化に繋がると考えられる。

既存の麻雀エージェントでは Convolutional Neural Network (CNN) を用いたモデルが多く用いられている。本研究では Transformer [3] を用いたアーキテクチャを提案する。提案手法では麻雀の盤面の情報を、各プレイヤーの行動が行われた順序を保持したままモデルに入力することが

¹ 東京大学工学部電子情報工学科
Department of Information and Communication Engineering, The University of Tokyo, Bunkyo, Tokyo 113-8654, Japan

² 東京大学工学部電気電子工学科
Department of Electrical and Electronic Engineering, The University of Tokyo, Bunkyo, Tokyo 113-8654, Japan

³ 東京大学大学院情報理工学系研究科電子情報学専攻
Department of Information and Communication Engineering, Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo, Tokyo 113-8654, Japan

a) ogami@logos.t.u-tokyo.ac.jp

b) tsuruoka@logos.t.u-tokyo.ac.jp

できる。そのため、捨て牌の前後関係などの順序が重要な情報を CNN を用いたモデルに比べてより正確に処理することができると考えられる。

2. 背景

2.1 麻雀の基本的なルールと用語

麻雀は、4人のプレイヤーがそれぞれ25000点の得点を持った状態でゲームが始まり、最終的に4人の中で最も高い得点を得ることを目指すゲームである。1試合は4または8局(局は試合の小単位)からなる。4人の中で最初に図1に示すような特定の和了形を作ったプレイヤーがその形に応じた得点を得ることで1局は終了する。

局において、各プレイヤーのターンで、プレイヤーは山から1枚牌を取る(ツモ)か、特定の条件を満たす場合に相手のプレイヤーが捨てた牌を取る(副露)を行うことができる。そして、和了を宣言するか、手牌の中から1枚捨てるかのどちらかを行う。

和了形が完成する場合として、自分が山から取った牌で和了形が完成する場合(ツモ和了)と、他プレイヤーが捨てた牌によって和了形が完成する場合(ロン和了)がある。特にロン和了の場合には、和了形を完成させたプレイヤーが得る得点の全てを、和了形を完成させる牌を捨てたプレイヤーが支払うことになる。このため、相手の和了形を完成させる牌を捨てないように行動することも重要である。

麻雀牌は34種類の牌4枚ずつの、合計136枚からなる。数字が書かれている牌は萬子、筒子、索子の3色に分かれており、それぞれの色について1から9の数字を表す牌がある。数字が書かれていない牌である字牌は7種類存在する。一般的なルールでは、そのうち数字の5を表す牌3種類それぞれについて、4枚中1枚が赤牌となっている。これを区別すると、図2に示すように牌の種類は37種類となる。

そのほかの用語について説明する。あるプレイヤーから見て、他3人のプレイヤーを、右側、正面、左側のプレイヤーをそれぞれ下家、対面、上家と呼ぶ。あと一枚で和了形が完成する状態のことを聴牌と呼び、和了形までの最短手数を向聴数と呼ぶ。

2.2 手牌推定の根拠

麻雀では、それぞれのプレイヤーが順番に山から1枚の牌を引き、自分の手牌の中から不要な牌を1枚捨てるという操作を繰り返す。その際、図1に示したような特定の和了形を目指してどの牌を捨てるかを決定する。和了形は主に連続する数字の牌3枚を用いて作られるため、「序盤に捨てられた牌と、その牌の周辺の牌を持っている可能性が低い」、「中盤以降に捨てられた牌の周辺の牌を持っている可能性が高い」といった推測を行うことが可能である。また、和了形は、少なくとも一つの役を満たしている必要が



図1 麻雀における和了形



図2 麻雀牌の種類と順序

あり、可能性のある役を絞り込むことによって相手の手牌を推定することが可能になる。

また、2.1項で述べたように、他プレイヤーの和了形を完成させる牌を捨てないことも重要である。他プレイヤーを警戒して捨てる牌を選んでいる場合には、特定の和了形を目指して選択しているわけではないため、先ほどの推定が成り立たなくなる。すなわち、相手の捨てた牌それぞれに対して、相手が他のプレイヤーをどの程度警戒して判断を行ったかという情報を加味して推定を行うことが必要になる。相手が他のプレイヤーを警戒していたかどうかを判断するためには、各時点までの他のプレイヤーの行動を把握する必要がある。

2.3 Transformer

Transformer [3] は機械翻訳を行うための手法として提案されたモデルであり、現在では自然言語処理のさまざまなタスクへの応用 [4] や画像認識への応用 [5] が行われるなど、幅広く用いられている。Transformer は系列データを入力として、系列データを出力することができるモデルである。Transformer の構成要素である注意機構では、入力された系列データのうち、任意の2点間の相互作用を直接計算することができるため、CNN に比べて長期の依存関係をより効率的に捉えることが可能である。



図3 手牌推測の正解ラベルの一例

3. 関連研究

既存の麻雀エージェントの多く [1], [2] はどの牌を捨てるかを決定する (打牌選択) モデルとして, CNN を主に用いたアーキテクチャを用いている. 打牌選択のモデルとして注意機構を用いた研究 [6] では盤面の情報を 373 次元のベクトルとして表し, マルチヘッドアテンションを用いた. 麻雀における不完全情報の推定の研究としては, 以下の研究が報告されている. 水上ら [7] は相手の手牌の向聴数, 聴牌の有無, 得点を線形回帰, ロジスティック回帰により推定した. 嶋村ら [8] は Bidirectional Long Short Term Memory (BiLSTM) をモデルとして用い, 役予測に必要な盤面情報について検討した.

本研究に最も近い研究は 根本らによる研究 [9] であり, 相手の手牌に 34 種類の牌がそれぞれ何枚ずつあるかを, 系列ラベリング問題を解くために用いられる手法である Conditional Random Field (CRF) [10] を用いて推定した. この研究で用いられた特徴量は, それぞれの牌の種類に関して周辺の牌が場に捨てられた枚数といった, 見えている牌の枚数に関するものだけであり, 牌が捨てられた順序などの, 手牌の推定に必要な特徴量が十分に組み込まれているとは言い難い. 相手 3 人の手牌を予測する際に各プレイヤーの手牌を個別に推定している点も本研究と異なっている. 各プレイヤーの手牌を個別に推定することで, それぞれの牌が最大で 4 枚しか存在しないという制約が守られないという懸念がある.

4. 提案手法

4.1 概要

麻雀におけるある 1 場面の情報を系列データとして表し, それを Transformer に入力し, 相手 3 人の手牌を系列データとして出力することにより相手の手牌を推定する.

モデルのアーキテクチャを図 5 に示す. モデルのエンコーダー側にはプレイヤーから観測可能な情報を特徴量へと変換して入力する. デコーダー側はエンコーダーから出力された特徴ベクトルを受け取り, 相手の手牌を自己回帰的に出力する.

4.2 相手の手牌の系列表現

図 3 の局面を例に説明する. 相手 3 人の手牌のうち, 推定を行うプレイヤーから見えていない部分 (図 3 において緑色の枠で囲まれた部分) を図 4 に示すように, 1 つの系列データとして表す. 相手 3 人の手牌を, 推定を行うプレイヤーから見た下家, 対面, 上家の順に並べ, プレイヤーの区切りに, プレイヤーの切り替わりを表す TRANS トークンを挿入する. 各プレイヤーの手牌は図 2 に示すように, 萬子の 1 から 9, 筒子の 1 から 9, 索子の 1 から 9, 字牌の順で並べる. それぞれ系列の開始と終了を表すクラスである BOS トークン, EOS トークンを系列の先頭, 末尾にそれぞれ設ける. 相手 3 人の手牌を別々に推定するのではなく, 1 つの系列として推定することによって, 牌の枚数に関する制約を満たした出力を得ることができると考えられる.

4.3 デコーダー

デコーダーは, エンコーダーが出力した特徴ベクトル f と出力系列の i 番目までを受け取り, $i+1$ 番目の系列が各トークンである確率 $P(y_{i+1} = k | y_0 \dots y_i, f) (k = 0, \dots, 39)$ を出力する. 訓練時には, f と出力系列の i 番目までを受け取り, 出力系列の $i+1$ 番目の確率が高くなるように学習を行う. 推論時には, 図 6 に示すように, f と出力系列の i 番目までを用いて $i+1$ 番目の確率分布を出力する. この確率分布を用いてビームサーチを行い, EOS が出力されるまで系列を生成することで, 相手 3 人分の手牌を推定することができる.

4.4 盤面の系列表現

推定を行うプレイヤーから見た盤面情報を系列データとして表現する手法について説明する. 盤面を表す情報としては, 誰がどのような牌を捨てているか, 親番が誰であるか, 各プレイヤーがリーチをしているか, などのさまざまな情報がある. これらの情報を表す際に, より小さな単位の特徴に分割し, それらの特徴ベクトルの合計として表す. 図 7 にその一例を示す. 推定を行うプレイヤーの上家のプレイヤーがリーチをしているかという情報を, プレイヤーから見た上家を表す特徴ベクトルと, リーチをしているかという特徴ベクトルの合計として表す. 表 1 にはここで用いた最小単位の特徴ベクトルの一覧を示す.

また, 麻雀の盤面を表す情報を大きく 2 つに分割する. 一方は各プレイヤーの行動の履歴を表す情報であり, 各プ



図 4 相手 3 人の手牌を系列データとして表現する方法

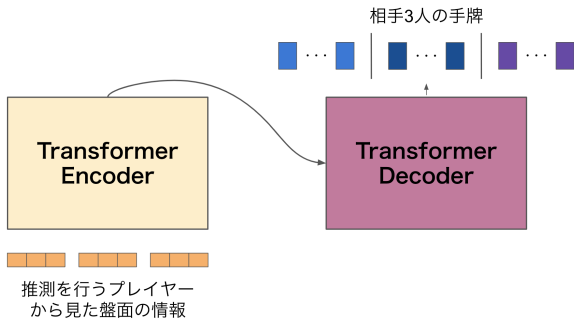


図 5 モデルの概略

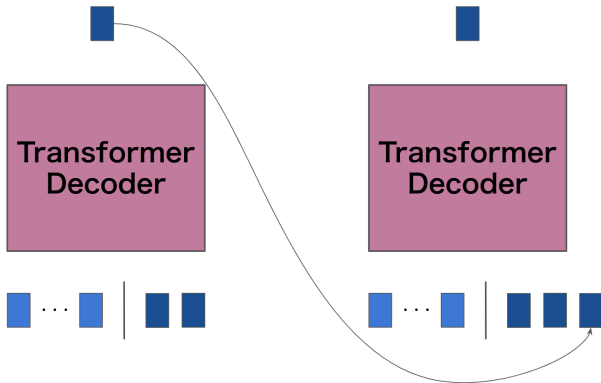


図 6 モデルのデコーダー部分

プレイヤーが自分の手牌から 1 枚捨てる行動や、副露、リーチ、新ドラの追加が該当する。もう一方は現在の状態を表す情報であり、場風や本場、手牌推定を行うプレイヤー自身の手牌などが該当する。図 8 に示すように、現在の状態を表す情報は系列の先頭から特定の順序で入力する。一方、時系列情報は実際に起こった順序で入力する。このようにすることで、プレイヤーの行動の詳細な前後関係の情報を失うことなくモデルに入力することができる。

実際に入力した盤面情報の表現方法およびその順序を表 2 に示す。表中に、それらの情報をどのように小単位の特徴ベクトルの和として表現したかについても示す。

5. 実験

5.1 平均一致枚数

麻雀において、相手の手牌のうち見えていない部分の枚

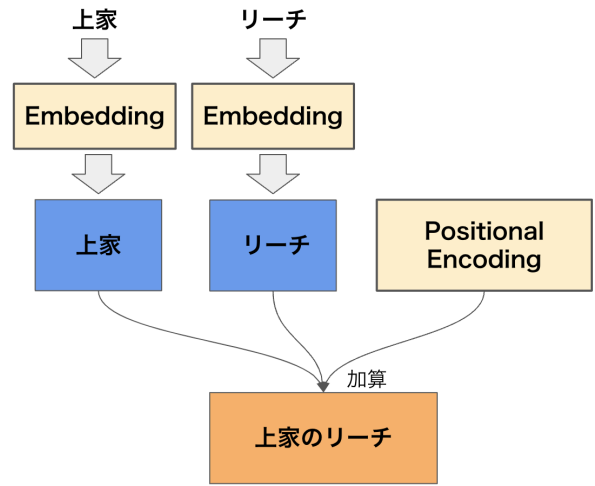


図 7 上家のリーチに関する特徴ベクトルの例

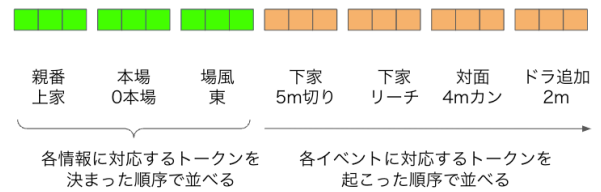


図 8 エンコーダー側の特徴量の順序

表 1 エンコーダー側の特徴ベクトルの最小単位

変数名	表す情報	トークン数
tile	牌	37
field_wind	場風	4
honba	本場	6
leader	親番	4
deposit	供託の本数	6
point	持ち点	10
tenpai	聴牌の有無	2
riichi	リーチをしているかどうか	2
double_riichi	ダブルリーチをしているかどうか	2
is_ippatsu	一発の状態かどうか	2
concealed	門前かどうか	2
player_hand	プレイヤーの手牌である	1
info_who	どのプレイヤーの情報か	4
field_info	場の情報である	1
dora_info	ドラの情報である	1
tedashi	手出しツモ切り	2
meld_by_who	誰が副露を行うか	4
meld_from_who	誰から鳴くか	4
meld_type	副露の種類	2

数は試合が進行するにつれて減少する。そのため、手牌推定の性能を測る指標として、平均一致枚数を用いた。平均一致枚数は相手の見えていない手牌 1 枚に対し、推定した手牌が何枚一致しているかを表す。手牌の推定を行うプレイヤーから見た相手プレイヤーを下家、対面、上家の順に 0, 1, 2 とする。麻雀には赤牌を区別すると 37 種類、区別しないと 34 種類の牌がある。ここで、牌の種類を 37 種類

表 2 盤面情報の系列化

情報	使用した変数	系列長
現在の状態を表す情報		
推定を行うプレイヤーの手牌	player_hand, tile	可変
場風	field_wind, field_info	1
親番	leader, field_info	1
供託の本数	deposit, field_info	1
得点	point, player_info	1
推定を行うプレイヤーが聴牌か否か	concealed, player_info	1
リーチしているかどうか	riichi, player_info	4
ダブルリーチしているかどうか	double_riichi, player_info	4
一発かどうか	ippatsu, player_info	4
門前かどうか	concealed, player_info	4
各プレイヤーの行動の履歴を表す情報		
ドラの追加	tile, dora_info	1
打牌	tile, player_info, tedashi	1
リーチ	riichi_event, player_info	1
副露	meld_by_who, meld_type, meld_from_who, tile	1 ~ 4

とした場合の手牌の枚数 $L(A)$, 手牌 $A = \{a_0, a_1, \dots, a_{36}\}$ と手牌 $B = \{b_0, b_1, \dots, b_{36}\}$ との一致枚数 $M(A, B)$, をそれぞれ以下のように定義する.

$$L(A) = \sum_{i=0}^{36} a_i \quad (1)$$

$$M(A, B) = \sum_{i=0}^{36} \min(a_i, b_i) \quad (2)$$

推定対象のプレイヤーの手牌をそれぞれ H_0, H_1, H_2 , 推定結果をそれぞれ R_0, R_1, R_2 とする. この時, 平均一致枚数 c は以下のように表される.

$$c = \frac{M(H_0, R_0) + M(H_1, R_1) + M(H_2, R_2)}{L(H_0) + L(H_1) + L(H_2)} \quad (3)$$

5.2 ランダム手法

提案手法との比較を行うために, 推定を行うプレイヤーの視点から見えていない牌の中から重複が起らないようにランダムに牌を選び, 手牌を生成する方法を用いた. 実験の結果, ランダムな手法の平均一致枚数は牌を 34 種類として計算すると 0.313 であり, 先行研究 [9] に示されたランダム手法の平均一致枚数と同様の結果が得られた. また, 赤牌と黒牌を区別し, 牌の種類を 37 種類として計算すると 0.302 であった.

5.3 データセット

訓練, テストに用いるデータセットとして, 天鳳^{*1}の最上位のフィールドである, 鳳凰卓の牌譜を用いた. 訓練には 2016 年の牌譜のうち, 2,633,128 局面のデータを用いた. テストには, 訓練データに用いてないデータのうち, 59,592 局面のデータを用いた.

*1 <https://tenhou.net/>

表 3 平均一致枚数

手法	赤牌を区別しない	赤牌を区別する
ランダム手法	0.313	0.302
CRF	0.417	なし
提案手法 (ビーム幅 1)	0.453	0.445
提案手法 (ビーム幅 20)	0.456	0.448

表 4 対象プレイヤーごとの平均一致枚数

対象プレイヤー	平均一致枚数
下家	0.449
対面	0.444
上家	0.441

6. 結果

6.1 概要

提案手法, CRF を用いた関連研究 [9] (CRF) およびランダム手法を比較した結果を表 3 に示す. 先行研究 [9] では平均一致枚数を計算する際に, 赤牌を区別していない. そこで, 結果の比較を行うためにランダム手法, 提案手法については赤牌を区別する場合としない場合の両方の結果を記載している. 提案手法がランダム手法の結果を上回っており, ビーム幅が 20 のとき, ビーム幅が 1 の場合の結果を上回っており, 適切なビーム幅でビームサーチを行うことにより, 確率最大の候補を選び続ける場合よりも性能が改善していることが確認できる.

6.2 プレイヤーから見えている枚数

麻雀においては, ゲームの進行と共に捨て牌や相手の副露など, 手牌推定を行うプレイヤーが見ることのできる牌が増加していき, 他プレイヤーの手牌をより正確に行うことができると考えられる.

図 9 に推定を行うプレイヤーから見える牌の枚数と, 平

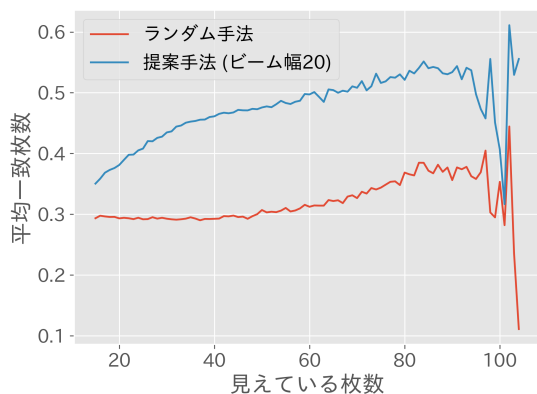


図 9 見えている枚数と平均一致枚数

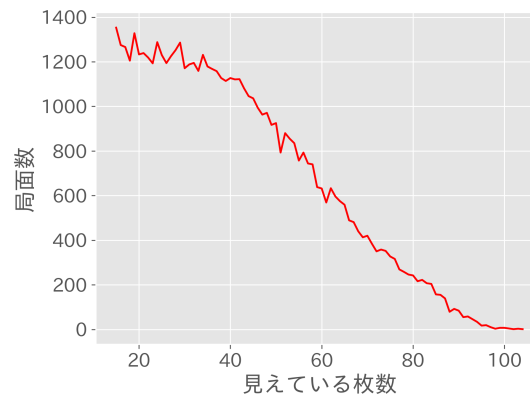


図 10 見えている枚数と局面数

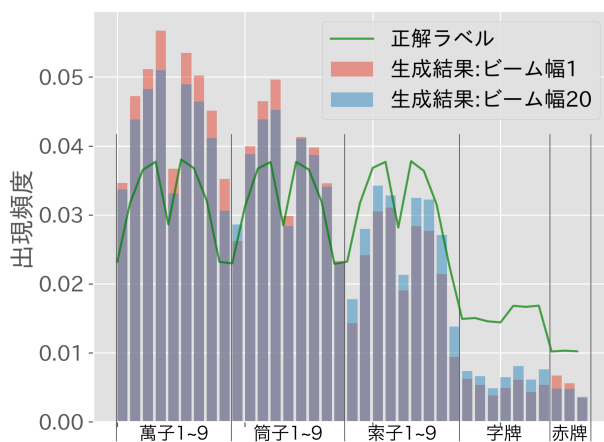


図 11 牌の種類ごとの出現頻度

均一致枚数との関係を表すグラフを示す。推定を行うプレイヤーから見える枚数の増加とともに、ランダム手法、提案手法のいずれも平均一致枚数が増加する傾向が見られた。提案手法の方が見えている枚数の増加に対する共に平均一致枚数の増加が急であり、プレイヤーから見える情報をより有効に利用できていると考えられる。図 10 に示すように見えている枚数が多い局面ほど数が少ないため、平均一致枚数のグラフの右端では平均一致枚数の値が不安定になっていると考えられる。

6.3 対象プレイヤーごとの平均一致枚数

対象プレイヤーごとに平均一致枚数を算出した結果を表 4 に示す。下家、対面、上家の順に平均一致枚数が低下している傾向が見られた。提案手法では、手牌を生成する際に、下家、対面、上家の順に手牌を生成している。手牌の生成が進めば進むほど、学習データとテストデータの乖離が大きくなるのが原因だと考えられる。

6.4 牌の種類ごとの出現頻度

提案手法が生成した手牌と、テストデータについて、牌の種類とその出現頻度の関係を図 11 に示す。テストデー

タでは萬子、筒子、索子の割合が均等である一方、提案手法が生成した手牌においては牌の出現頻度に偏りがある。萬子、筒子の割合が索子に比べて多くなっており、字牌や赤牌の生成確率は小さくなっている。提案手法において図 2 に示した順序で手牌を生成していることが原因だと考えられる。ビーム幅が 20 の時、ビーム幅が 1 の時に比べて偏りが小さくなっており、ビームサーチを行うことで、手牌を系列として表す際に用いる牌の順序に依存した偏りを軽減することができると言える。

7. 考察

提案手法では、麻雀の盤面情報を系列データとして表し、Transformer に入力することによって相手の手牌を推定した。その結果、ランダム手法および関連研究 [9] よりも高い平均一致枚数が得られた。

一方で、6.4 項に示したような、訓練データにおける牌の順序に起因すると考えられる牌の偏りが生じた。このように、平均一致枚数が改善されたとしても望ましくないような推定結果を出力してしまうことがあるため、平均一致枚数以外の評価指標もあわせて検討することが必要である。本研究では一般的に麻雀牌を表示する際に用いられる順序で手牌を並べた。牌の種類ごとに枚数の推定の難易度が異なるため、より確実に枚数を推定できると考えられるような牌を前半に並べることにより、訓練データとテストデータにおける分布の乖離を抑制しやすくなると考えられる。

手牌推定において、見えている牌の枚数や相手の捨て牌の違いにより、推定の確信度合いが異なる。図 10 に示したように、今回用いたデータセットの中には推定を行うプレイヤーから見えている牌の枚数が少ない局面など、推定が難しい局面も多く含まれている。そのような局面のサンプル数を調整することにより、本来高い確信度で推定を行うことができるような局面での精度を向上させることができると考えられる。

8. おわりに

本稿では、麻雀の盤面情報を系列データとして表現する手法を提案し、系列データを処理することができるモデルである Transformer を用いて相手の手牌を自己回帰的に推定した。その結果、CRF を用いて手牌推定を行なった研究 [9] や、プレイヤーから見えている牌を考慮することでランダムに手牌を推定する手法に比べて高い平均一致枚数を出すことができた。

一方で、ビームサーチを行なった場合でも生成される手牌における牌の種類に、訓練データにおける牌の並び順に起因する偏りが見られた。今後の課題としては、前述した牌の種類偏りを改善することや、より手牌推定を行いやすいようなデータに絞って訓練を行うことが考えられる。

参考文献

- [1] Junjie Li, Sotetsu Koyamada, Qiwei Ye, Guoqing Liu, Chao Wang, Ruihan Yang, Li Zhao, Tao Qin, Tie-Yan Liu, and Hsiao-Wuen Hon. Suphx: Mastering mahjong with deep reinforcement learning. *arXiv preprint arXiv:2003.13590*, 2020.
- [2] Naga. https://dmv.nico/ja/articles/mahjong_ai_naga/. (Accessed on 09/30/2022).
- [3] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, Vol. 30. Curran Associates, Inc., 2017.
- [4] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, Vol. 21, No. 140, pp. 1–67, 2020.
- [5] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021.
- [6] Trong Duc Truong. A supervised attention-based multiclass classifier for tile discarding in japanese mahjong. Master's thesis, University of Agder, 2021.
- [7] 水上直紀, 鶴岡慶雅ほか. 牌譜を用いた対戦相手のモデル化とモンテカルロ法によるコンピュータ麻雀プレイヤーの構築. ゲームプログラミングワークショップ 2014 論文集, Vol. 2014, pp. 48–55, 2014.
- [8] 嶋村文里, 松澤智史ほか. 予測精度の変化を用いた麻雀の役解析. 研究報告ゲーム情報学 (GI), Vol. 2022, No. 9, pp. 1–5, 2022.
- [9] 根本佳典, 古宮嘉那子, 小谷善行ほか. Crf を用いた麻雀の不完全情報推定. ゲームプログラミングワークショップ 2012 論文集, Vol. 2012, No. 6, pp. 155–158, 2012.
- [10] John D. Lafferty, Andrew McCallum, and Fernando C. N. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, p. 282–289, San Francisco, CA,

USA, 2001. Morgan Kaufmann Publishers Inc.

- [11] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.
- [12] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

付 録

A 実験の詳細

表 A.1 にモデルのパラメータを示す。最適化には Adam [11] を使用し、パラメータは $\beta_1 = 0.9$, $\beta_2 = 0.999$ とした。学習率は開始から 100 ステップの間は線形に増加させ、最大値 0.009 をとるようにした。その後は 162,356 ステップまでの間、ステップ数の平方根の逆数に比例して減少させた。正解クラスの確率を 1 よりも小さい値にし、正解でないクラスの確率を 0 より大きい値にするラベル平滑化 [12] を行った。

表 A.1 モデルのパラメータ

モデルのパラメータ	値
エンコーダー側の層数	4
デコーダー側の層数	4
アテンションヘッドの数	8
活性化関数	ReLU
特徴量ベクトルの次元数	512
フィードフォワードネットワークの次元数	2048
ドロップアウト	0.1

B 手牌推定の実例

訓練データに含まれていない牌譜*2を用いて、実際に手牌推定を行なった結果の一例を示す。図 B.1 の局面で手牌推定を行ったところ、図 B.2 に示すような結果が得られた。副露を行なっている下家、上家の手牌は聴牌している形だと推定している。一方で、副露を行なっている他のプレイヤーに対し、リスクの低い選択をしているように見える対面の手牌は聴牌から遠い形になっていると推定している。

C 副露に関する情報の系列化

副露にはポン、チー、大明槓、暗槓、加槓の 5 種類がある。このうち、暗槓、加槓は自分が山から引いてきた牌を使って行う行動であり、必ず同じ種類の牌を 4 枚用いる。ポン、チー、大明槓は、他のプレイヤーが捨てた牌を使って行う行動である。

暗槓、加槓の場合は、表 1 における、meld_by_who、

*2 <https://tenhou.net/3/?log=2020120117gm-00a9-0000-d737b979tw=3>



図 B.1 手牌推定を行った局面

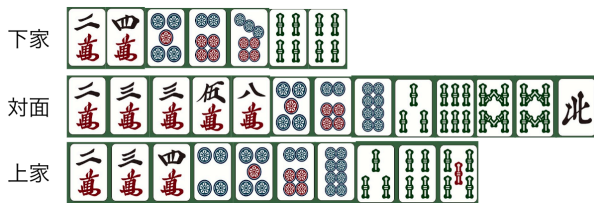


図 B.2 手牌推定の結果

meld_type, tile の表現ベクトルの和として表す。加槓の場合には実際に加えた牌を tile を用いて表現することにより、赤牌を加えたか否かという情報を保持することができる。暗槓と加槓の場合、系列長は 1 となる。

ポン、チー、大明槓の場合は副露に用いる牌ごとに長さ 1 の系列となるように、meld_by_who, meld_from_who, tile を用いて表す。この時、系列長はポン、チーの場合には 3、大明槓の場合には 4 になる。図 C.3 のように、副露に用いる牌のうち、相手の捨て牌からとった牌を先頭に、副露に用いた牌を表す tile ベクトルを並べ、そのそれぞれに meld_by_who, meld_from_who ベクトルを加算する。

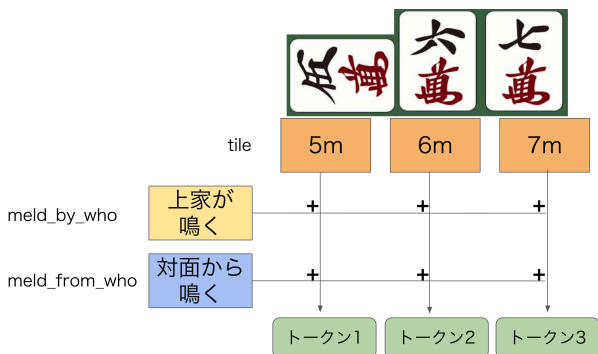


図 C.3 ポン、チー、大明槓の表現方法