

# 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

水上直紀<sup>1</sup> 鶴岡慶雅<sup>1</sup>

<sup>1</sup>東京大学

# 背景

- コンピュータポーカーは世界チャンピオンレベル  
(2人) [Bowling+ 2009]
  - 多人数ポーカーの搾取 [Van 2010]
    - モデル化:棋譜から相手のカードと行動の推定
    - モンテカルロ木探索
- モデル化とシミュレーションを麻雀にも適用

# これまでの研究 [水上+ 2013]

- 一人麻雀との差の解析による四人麻雀プレイヤーの作成
- 一人麻雀を定義
  - 鳴きはできない
  - 和了率のみで評価
- 降りる局面の学習
  - 牌譜中の手に対して、降りかどうかのタグ付を行う
  - 現局面が降りるべきかどうかの2値問題として学習

# これまでの研究 [水上+ 2014]

- 鳴きについての学習
  - 一人麻雀の手の学習方法を鳴きに拡張  
(以後、一人麻雀の手と呼ぶ)
  - F1値 0.86 と高精度
- 平均プレイヤーよりも強い実力
- 教師データを大量に確保するのが困難
- 回し打ちが出来ない

# 提案手法

- 牌譜を用いた相手のモデルの抽象化とモンテカルロ法による麻雀プレイヤーの作成
- 相手モデルの抽象化
  - 聴牌しているか
  - 待ち牌は何か
  - 和了点はいくつか
- 利点
  - 相手の手牌の予測が不要
  - 牌譜のみから学習が可能

# 提案手法

- 牌譜を用いた相手のモデルの抽象化とモンテカルロ法による麻雀プレイヤーの作成
- モンテカルロ法による手の決定
  - モデル化した相手と一人麻雀の手によるシミュレーション
- 利点
  - 自分と相手の手牌の良さを同じ尺度で比較可能
  - すべての手を考慮に入れることができる

# 研究の概要

- 相手プレイヤーの聴牌予測
- 待ち牌の予測
- 得点の予測
- モデル化とモンテカルロ法を用いた手の決定
- 結果

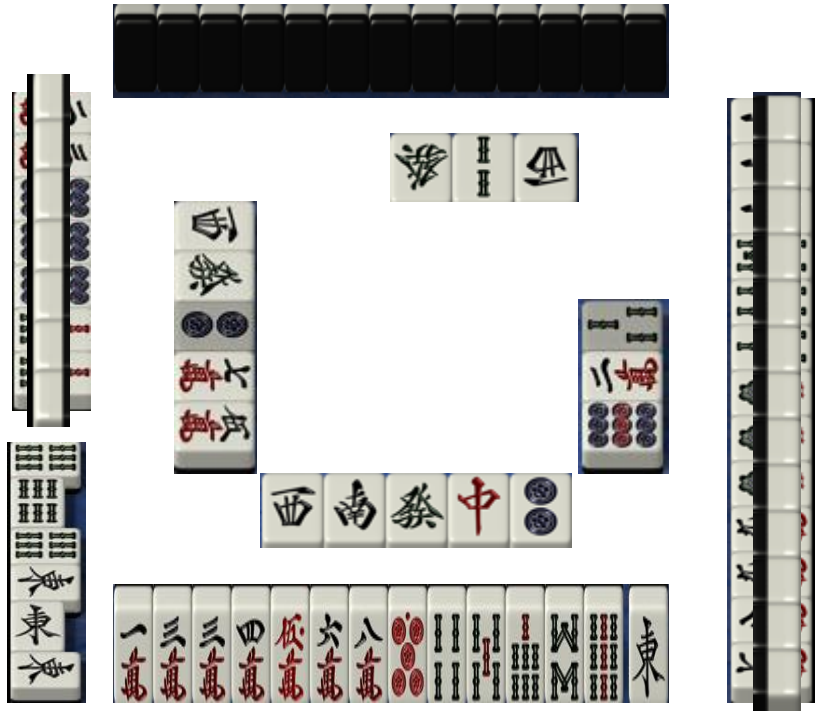
# 研究の概要

- 相手プレイヤーの聴牌予測
- 待ち牌の予測
- 得点の予測
- モデル化とモンテカルロ法を用いた手の決定
- 結果



# 聴牌予測とは

- あるプレイヤーの視点から別のプレイヤーが聴牌しているかどうかを予測する

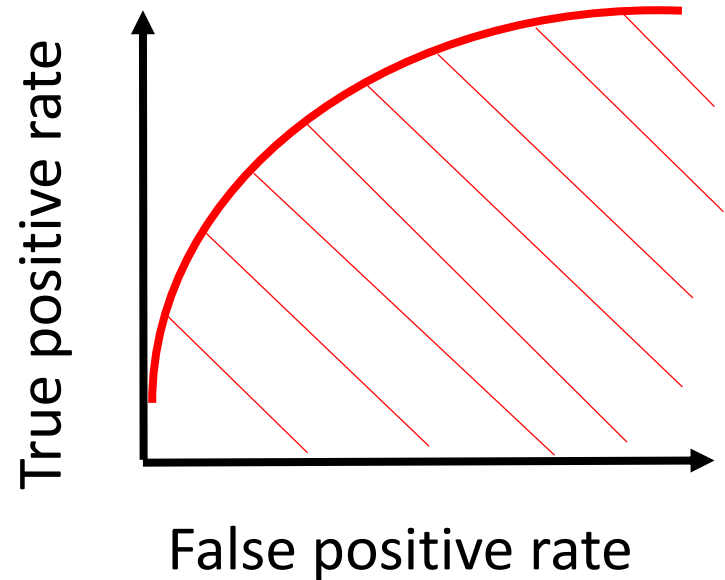


# 学習設定

- データセット
  - 教師データ
    - $1.77 \times 10^7$ 局面
  - テストデータ
    - 100局面
    - リーチしたプレイヤーはデータにない
- 特徴量
  - リーチ、副露数、順目等
  - 6,888次元
- ロジスティック回帰モデル

# 評価方法と結果

- 評価方法
  - ROC 曲線下面積 (AUC)



プレイヤー	AUC
上級者	0.778
分類器	0.777

上級者と同等の結果

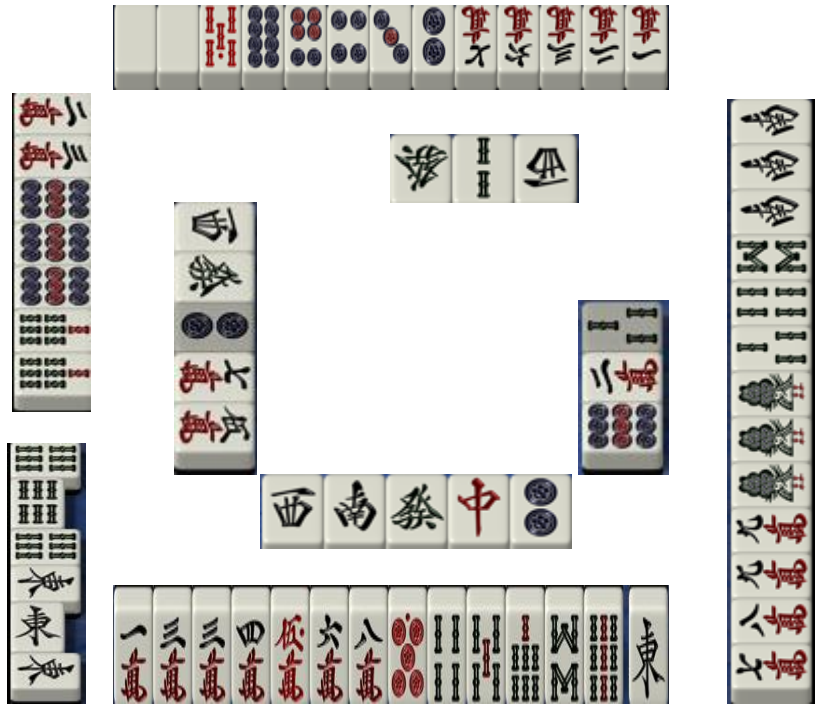
上級者: 上位0.1%のプレイヤー

# 研究の概要

- 相手プレイヤーの聴牌予測
- 待ち牌の予測
- 得点の予測
- モデル化とモンテカルロ法を用いた手の決定
- 結果

# 待ち牌予測とは

- プレイヤーが聴牌しているか既知である場合に相手の待ち牌を予測する



# 学習設定

- データセット
  - 教師データ
    - $7.24 \times 10^7$ 局面
  - テストデータ
    - 100局面
- 特徴量
  - 捨て牌、副露の種類、見えている牌の枚数等
  - 31,416次元
- ロジスティック回帰モデル

# 評価方法

相手の手



自分の手

1 2 ...



$$\text{スコア} = 4 / (12-2)$$

# 結果

- ランダム:手牌をランダムに切る

プレイヤー	スコア
上級者	0.744
分類器	0.676
ランダム	0.502

上級者には劣る



# 研究の概要

- 相手プレイヤーの聴牌予測
- 待ち牌の予測
- **得点の予測**
- モデル化とモンテカルロ法を用いた手の決定
- 結果

# 得点予測とは

- プレイヤに対して和了牌をロンされた場合に支払う得点の予測
- 高い翻数のデータの影響力が大きい  
→得点の自然対数をとった値を教師データとする



# 学習設定

- データセット
  - 教師データ
    - $5.92 \times 10^7$ 局面
  - テストデータ
    - 100局面
- 特徴量
  - リーチ、見えていないドラ、副露の種類等
  - 26,889次元
- 重回帰モデル

# 評価方法と結果

- 平均二乗誤差(Mean Squared Error)を用いる

プレイヤー	MSE
上級者	0.40
分類器	0.37

上級者より**高精度**

# 研究の概要

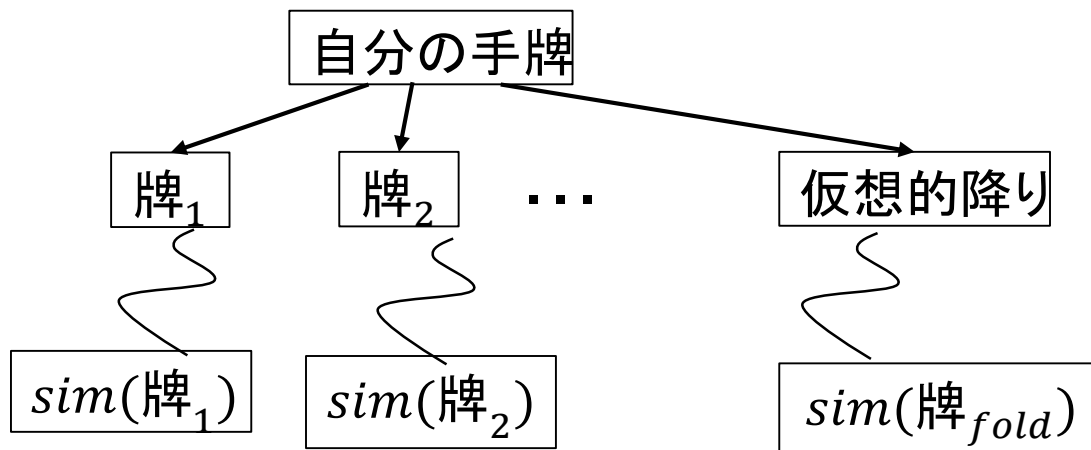
- 相手プレイヤーの聴牌予測
- 待ち牌の予測
- 得点の予測
- **モデル化とモンテカルロ法を用いた  
手の決定**
- 結果

# モンテカルロ法の手決定方法

手の決定には二つの項の計算が必要

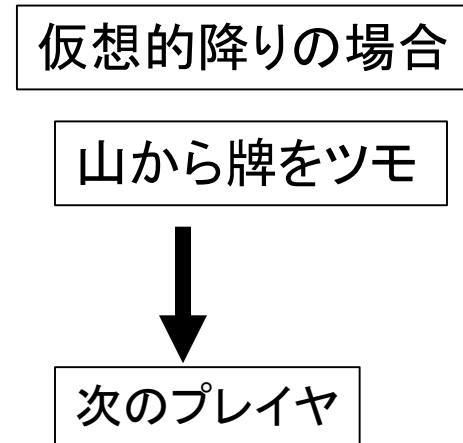
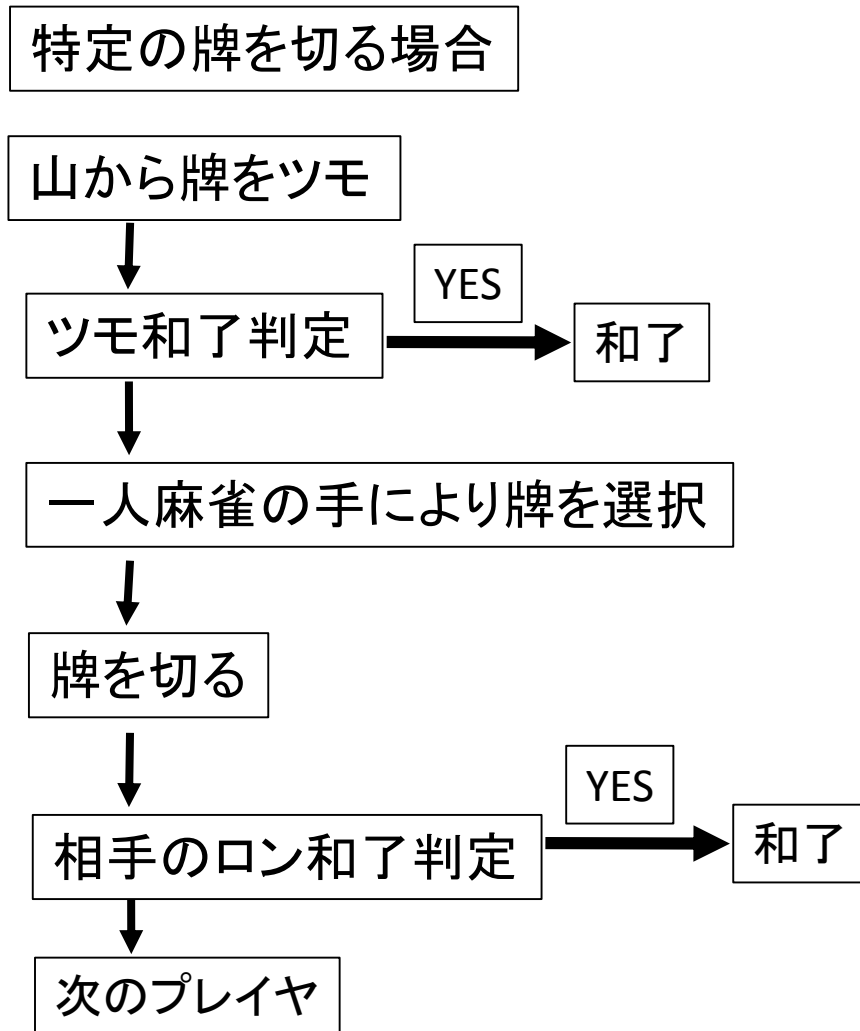
$$\text{Score}(\text{牌}) = \text{Max}(\text{sim}(\text{牌}), \text{sim}(\text{牌}_{\text{fold}})) - \sum_{\text{他家}=1}^3 \text{EV}(\text{他家}, \text{牌})$$

$$\text{EV}(\text{他家}, \text{牌}) = \text{Pr}(\text{聴牌}(\text{他家}) = \text{YES}, \text{牌} = \text{待ち牌}) * \text{得点}(\text{他家}, \text{牌})$$

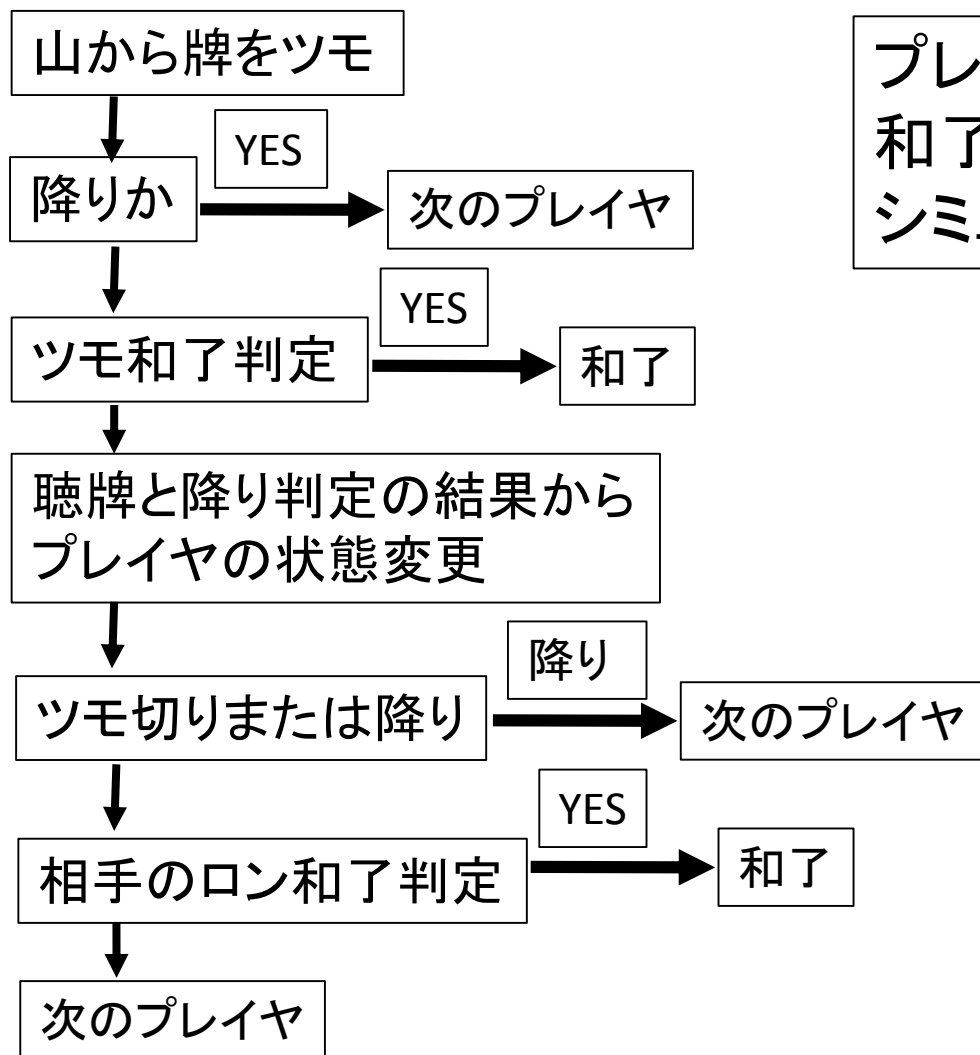


降り続けた値として  
 $\text{sim}(\text{牌}_{\text{fold}})$ を計算

# シミュレーション中の自分の手番



# シミュレーション中の相手の手番

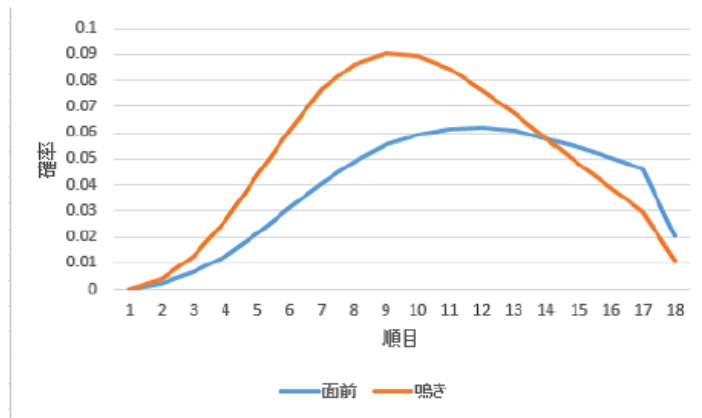


プレイヤーの状態（聴牌と降り）と和了牌の確率分布をもとにシミュレーションを行う。



# プレイヤーの聴牌状態の変更

- 一人麻雀の手を用いて各順目に聴牌する確率を計算



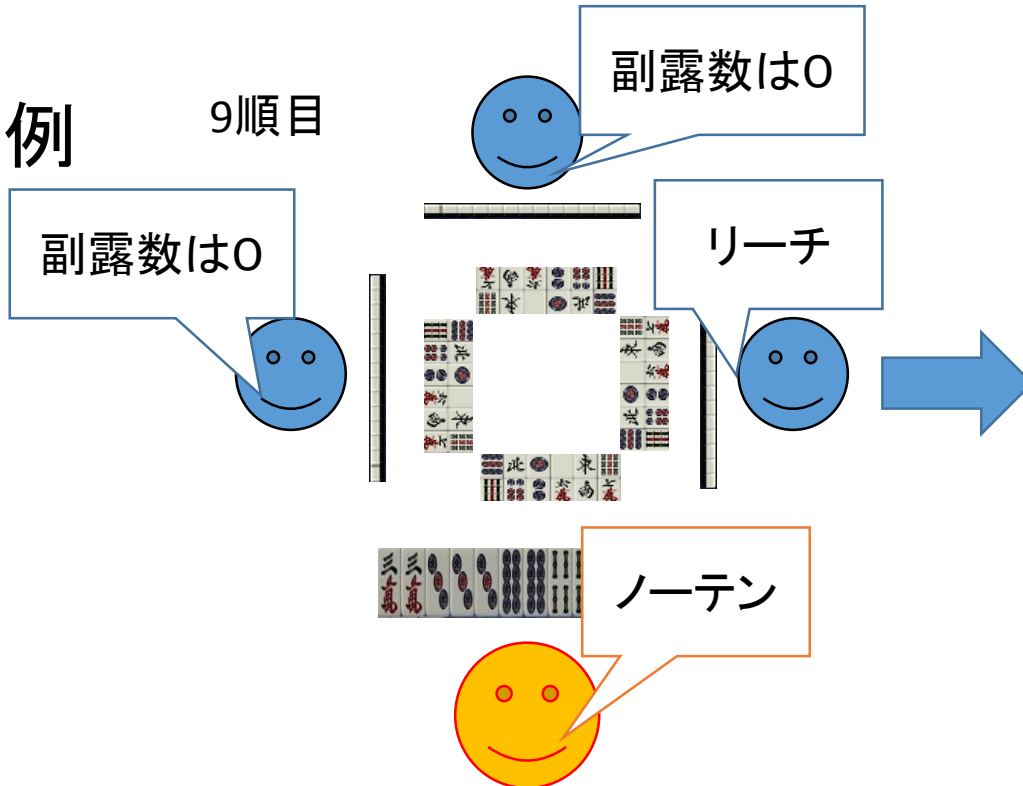
- シミュレーション開始時に面前か鳴きか決定
- 乱数を用いてプレイヤー状態を変更する

# プレイヤーの降り状態の変更

- 牌譜中の場の状況と一人麻雀の手的一致率から降りる確率を求める

例

9順目



この時の牌譜と一人麻雀の手的一致率は60%

この場合、40%で降りとする

# 手の決定

- 序盤ではモンテカルロ法の手は悪いと予想  
→モンテカルロ法の手と一人麻雀の手の使い分け
- $$\begin{cases} \text{一人麻雀の手} & \text{if } (\text{Pr}(\text{聴牌}(\text{他家}) = \text{YES}) < \alpha) \\ \text{モンテカルロ法の手} & \text{otherwise} \end{cases}$$
- 閾値 $\alpha$ をどう決定するか  
→牌譜との一致率

# 手の決定

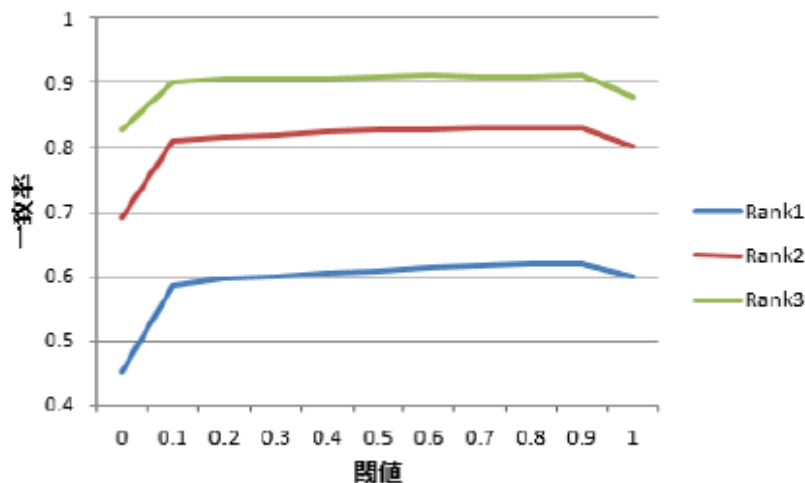
- 設定

- テストデータ:  $10^5$ 局面

- 評価

- Rank  $n$ は第 $n$ 候補に牌譜での打牌が含む割合

- 閾値は0から1まで0.1刻み



閾値  $\alpha = 0.9$  が最も一致率が高い

# 研究の概要

- 相手プレイヤーの聴牌予測
- 待ち牌の予測
- 得点の予測
- モデル化とモンテカルロ法を用いた手の決定
- **結果**

# 評価

- 提案手法と[水上+ 2014]の比較
- 持ち時間 一手一秒
- 赤あり東風戦
  
- 対コンピュータ麻雀
  - まったり麻雀
  - デュプリケートモード 初期乱数0~999の1000局
- 対人戦
  - インターネット麻雀サイト“天鳳”の上級卓

# まったり麻雀

- 強さ

- 発表者の知る限り最も強い（発表者と同等の強さ?）
- インターネット麻雀サイトでのレーティングはない

- 評価関数

- 統計量をヒューリスティックに組み合わせたもの
- HPに簡単な説明
- ソースコードは非公開

# まったり麻雀

VS まったり麻雀	1位率	2位率	3位率	4位率	平均順位
まったり麻雀	0.248	0.247	0.250	0.255	2.51±0.03
提案手法	0.230	0.248	0.268	0.254	2.54±0.03
[水上+ 2014]	0.243	0.226	0.222	0.309	2.59±0.03

プレイヤー	和了率	放銃率
まったり麻雀	0.200	0.122
提案手法	0.201	0.121
[水上+ 2014]	0.228	0.178

勝ち越すことはできなかったが  
成績が**向上**した



# 対人戦

	1位率	2位率	3位率	4位率	平均順位	試合数
提案手法	0.231	0.269	0.268	0.233	2.50±0.02	2227
[水上+ 2014]	0.253	0.248	0.251	0.248	2.49±0.03	1441

プレイヤー	安定レーティング	保障安定レーティング
人間トップ	N/A	2150
提案手法	1681	1690
[水上+ 2014]	1689	1610
平均プレイヤー	N/A	1500

保障安定レーティング  
が**向上**した

プレイヤー	和了率	放銃率
提案手法	0.237	0.127
[水上+ 2014]	0.256	0.148

# まとめ

- モデル化の各要素の精度は比較的高い
- モンテカルロ法を行うことで、今以上の実力をもった麻雀プレイヤーを得た
  
- 得点を考慮した手の選択
- シミュレーションの改善
- 一人麻雀の手の改善

# 特徴量

- リーチを打っているか
- 副露数と捨て牌の数
- 副露数とその順目
- 副露数と手出しの数
- 副露数と最後に手出した牌の種類
- 副露した種類と副露した時に切った牌の種類
- ドラの種類を切ったか
- 赤ドラを切ったか
- 手出し牌とその次の手出し牌の組み合わせ

特徴ベクトルの次元は全部で6,888

# 特徴量

- 牌の種類とその牌が何枚見えているか
- 現物かどうか
- $n$  回 ( $n=0, 1, 2$ ) 手出しの間に通った牌
- ドラの種類
- 順目とその時に捨てた牌
- リーチ時に切った牌
- 副露の種類と副露時に切った牌
- 二つの副露の種類
- 捨て牌二つの種類と手出しかどうか
- 手出し牌とその次の手出し牌の組み合わせ

特徴ベクトルの次元は全部で31,416

# 特徴量

- 親かどうかとリーチしているかどうか
- 確定している役と見えているドラの枚数
- リーチかダマか副露と確定している役と見えているドラの枚数
- 副露の種類と確定している役と見えているドラの枚数
- 二つの副露の種類
- 副露数と確定している役と見えているドラの枚数
- リーチしているかどうか切った牌が筋になっているかタンヤオ牌かどうか
- オタ風を鳴いた, さらに役牌を鳴いた, 何も鳴いていない
- 切った牌がドラ, ドラの一つ隣, 二つ隣, 同じ色, 無関係
- タンヤオが可能な副露とドラがタンヤオと見えてるドラの枚数
- ホンイツが可能な副露と確定している役とドラが染め色かどうか
- チンイツが可能な副露と確定している役とドラが染め色かどうか
- トイトイが可能な副露と確定している役とドラが染め色かどうか
- 三元牌が何種類鳴かれている
- 風牌が何種類鳴かれているか

特徴ベクトルの次元は全部で26,889