

CRF を用いた麻雀の不完全情報推定

根本佳典[†] 古宮嘉那子^{††} 小谷善行^{††}

麻雀は不完全情報ゲームの一つである。麻雀における見えない情報は、山と相手の手牌である。相手の手牌を推定することは、相手の待ち牌を読んだり、得点の期待値を計算する際に、有用であると考えられる。本論文では、系列ラベリング問題のための識別モデルである Conditional Random Fields(CRF) の着想を用いて、相手の手牌を推定することを提案する。評価方法として、テストデータとのラベルの一致率、実際の手牌と推定した手牌のシャンテン数を比較した。提案手法の手牌の一致率は 42% となった。

Estimation of Imperfect Information using CRF in Mahjong

YOSHINORI NEMOTO,[†] KANAKO KOMIYA^{††}
and YOSHIYUKI KOTANI^{††}

Mahjong is one of imperfect information games. Invisible information in Mahjong is the wall and opponents' hands. Predicting opponents' hands is useful to assume their waiting tiles and to calculate the expected value of the score. In this paper, we propose to estimate opponents' hands using Conditional Random Fields(CRF), which is a recognition model for sequence labeling problem. We compared the agreement rate between the estimated hands and the real hand in a game and the number of steps to the ready hand to win. An agreement rate of the hands using proposal method was 42%.

1. はじめに

麻雀には山や相手の手牌といった見えない情報があるため、不完全情報ゲームに分類される。麻雀の先行研究としては、UCT 探索を適用した UCT 探索を用いたもの¹⁾ などがある。

不完全情報である相手の手牌を推定することは、相手の必要としている牌や待ち牌を読んだり、得点の期待値を計算する際に、有用であると考えられる。また、相手の手牌は、捨て牌や場に見えている枚数などから推定される。

本研究では、系列ラベリング問題のための識別モデルである CRF(Conditional Random Fields) の着想を用いて、相手の手牌を推定することを提案する。今回は、入力として河や場に見えている牌の情報を与えて、34 種類の各牌に対して相手の手牌にある枚数をラベルとして出力する。出力されたラベルの枚数に対応した手牌を、相手の手牌として用いることで、相手の手牌の推定を行う。

2. CRF

2.1 CRF

Conditional Random Fields (CRF) とは、Lafferty らによって提案された、系列ラベリング問題に対数線形モデルに適用したものである²⁾³⁾。

$$D = \{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), (\mathbf{x}^{(2)}, \mathbf{y}^{(2)}), \dots, (\mathbf{x}^{(|D|)}, \mathbf{y}^{(|D|)})\}.$$

のようなデータが与えられたとき、その条件付き確率 $P(\mathbf{y}|\mathbf{x})$ は

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z_{\mathbf{x}, \mathbf{w}}} \exp(\mathbf{w} \cdot \phi(\mathbf{x}, \mathbf{y})) \quad (1)$$

$$Z_{\mathbf{x}, \mathbf{w}} = \sum_{\mathbf{y}} \exp(\mathbf{w} \cdot \phi(\mathbf{x}, \mathbf{y})) \quad (2)$$

- \mathbf{x} : 系列
- \mathbf{y} : 系列に付与されるラベル列
- \mathbf{w} : 重みベクトル
- $\phi(\mathbf{x}, \mathbf{y})$: 素性ベクトル
で表される。

入力 \mathbf{x} に対する最適な出力 \mathbf{y} は、Viterbi アルゴリズムを用いて効率的に求めることができる。

2.2 CRF++

CRF の実装として、オープンソースの CRF ツールキットである CRF++ を使用した⁴⁾。

[†] 東京農工大学大学院 工学府 情報工学専攻
Department of Computer and Information Sciences Graduate School of Engineering

^{††} 東京農工大学 工学研究院 先端情報科学部門
Institute of Engineering Tokyo University of Agriculture and Technology

3. 不完全情報の推定

3.1 CRF の着想を用いた不完全情報の推定

相手プレイヤーがある牌を持っていると予想とした場合、その周囲の牌も相手プレイヤーは持っていることが多い。このように、ある牌の枚数が、その周囲の牌の状態から影響をうけることから、相手プレイヤーの手牌推定は系列ラベリング問題の考え方が適用できるのではないかと考えた。系列ラベリング問題とは、言語処理の単語の品詞タグ付など、ある系列の各要素に適切なラベルを付与し、ラベル列を出力する問題を示す。

麻雀において、相手プレイヤーの手牌を予想する際に手がかりとなるのは、相手の河、鳴き面子、場に見えている牌などである。牌の種類を一つの系列とみなし、牌の情報として予想する相手プレイヤーの河にある各牌の枚数と場に見えている各牌の枚数をラベルとして与える。

学習・推定用に使用したデータの一部を表 1 に示す。データは牌の種類、推定する相手の河にある枚数、場に見えている枚数、推定する相手の手牌にある枚数の 4 列からなる。牌の順番は萬子の 1 から 9、筒子の 1 から 9、索子の 1 から 9、東南西北發中とする。

システムは、牌の種類、推定する相手の河にある枚数、場に見えている枚数を入力とし、推定する相手の手牌にある枚数を学習する。推定では、相手の手牌にある枚数を出力し、その枚数に対応した手牌を相手の手牌として生成する。

表 1 実験用局面データ

牌	河	場	手牌	牌	河	場	手牌
1m	1	2	0	9p	0	1	0
2m	0	1	0	1s	0	3	0
3m	0	3	0	2s	0	2	2
4m	0	2	1	3s	0	1	0
5m	0	0	1	4s	0	1	0
6m	0	1	0	5s	0	1	0
7m	0	0	3	6s	0	0	0
8m	0	1	1	7s	0	1	1
9m	1	1	0	8s	0	1	0
1p	2	0	0	9s	0	1	0
2p	0	0	0	東	0	1	0
3p	0	0	1	南	2	0	0
4p	0	0	1	西	0	1	0
5p	0	1	0	北	0	4	0
6p	0	0	0	白	3	1	0
7p	0	0	1	發	0	0	0
8p	0	0	1	中	1	1	0

3.2 CRF で用いた特徴量

今回採用した素性は

- 牌の種類
- 河にある枚数
- 2 つ連続した牌の河にある枚数
- 場に見えている枚数

- 2 つ連続した牌の場に見えている枚数
- 3 つ連続した牌の河にある枚数
- 3 つ連続した牌の場に見えている枚数
- 牌の種類と河にある枚数
- 牌の種類と河にある枚数と 2 つ隣までの牌の場に見えている枚数

とした。

4. CRF を用いた不完全情報推定

4.1 CRF を用いた不完全情報推定の実験概要

CRF のツールキットである CRF++ を使用して、相手の手牌の推定を行う。

実験に使用した局面の数と素性の組み合わせの数を表 2 に示す。抽出した局面から、

- 牌種
- 推定する相手の河にある枚数
- 場に見えている枚数
- 推定する相手の手牌にある枚数

を調べ、それらを 34 種類の牌分並べたデータを作成する。また、素性の数とは、素性テンプレートと与えた局面のデータによって生成された素性の組合せの総数を指す。

この節では、局面の数と素性の数が不完全情報推定に与える影響を調査する。学習データは局面の数を 761 から 12101 まで変化させた。局面の数は大きく変化しているが、素性の数は局面の数の差に比べるとあまり違いが見られない。これは入力データの種類の種類が牌の種類と枚数を表すラベルのみで構成されることによる。素性として隣り合った牌のデータの組合せを採用しているが、牌の種類 34 とラベルの種類 5 または 6 の内から 2 または 3 個の要素の組合せとなるため、素性の組み合わせには限りがある。また、頻出するデータの組合せも多く存在する。これらの理由から、一定の数の局面数を超えていれば、それ以上局面の数を大きくしても、素性の数はあまり増加しない。

表 2 CRF を用いた不完全情報推定の学習・テストデータの局面の数と素性の数

データ名	局面	素性
1	761	21895
2	1157	23410
3	3081	25915
4	6205	27540
5	12101	29485
test	28740	30795

4.2 CRF を用いた不完全情報推定の実験結果

表 3 に open データのラベルの一致率を示す。

図 1 に CRF を用いた不完全情報推定の一致率の推移を示す。

CRF を用いた不完全情報推定では、ランダムに生成した手牌の一致率が 31.3% に対し、CRF を用いた不完全情報推定は最大 7.22% と大きく下回っている。

また、学習の局面数を大きくすればするほど、一致率が小さくなっている。

表 3 CRF を用いた不完全情報推定の一一致率

データ名	総数	一致数	一致率 [%]
1	3611149	26090	7.22417
2	3611149	15488	4.28853
3	3611149	5760	1.59491
3	3611149	4891	1.35429
5	3611149	4664	1.29143
ランダム	3611149	112892	31.2591

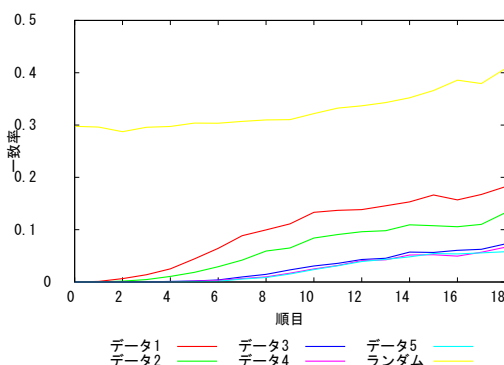


図 1 CRF を用いた不完全情報推定の一一致率の推移

表 4 に不完全情報推定システムが open データに対して付与したラベルの内訳を示す。

表 4 を見ると、すべてのデータがラベル 0 を最も多くつけていることがわかる。ラベル 0 は手牌にある牌が 0 枚ある、つまりある牌を持っていないと推定していることを表す。

この傾向は、局面の数と素性の数が大きくなるにつれて顕著になっている。test データのラベルの内訳と比較すると、不完全情報推定システムの予想したラベル 1 以上の数はかなり小さいことが分かる。

CRF を用いた不完全情報推定で出力されたラベルを元に生成した手牌は、枚数が 13 枚に届くことはなく、ほとんどの手牌が 0 枚となってしまっていた。この結果から、CRF をそのまま用いて不完全情報推定することは非常に難しいと言える。

4.3 ダミーを用いた学習

学習データは、萬子、筒子、索子、字牌の順に連続で並べていた。しかし、このままでは、九萬と一筒など連続で持っても面子を構成できない牌の並びや、字牌の並びも他の中張牌と同じように学習されてしまう。

この問題を解決するために、萬子、筒子、索子の間と各

表 4 CRF を用いた不完全情報推定のラベル内訳

データ名	学習局面数	ラベル 0	ラベル 1
1	761	926138	45198
2	1157	948179	26440
3	3081	967569	9410
4	6205	969124	7998
5	12101	969400	7738
テストデータ	687498	226080	55837

データ名	ラベル 2	ラベル 3	ラベル 4
1	5319	505	0
2	2328	195	18
3	181	0	0
4	38	0	0
5	22	0	0
テストデータ	55837	7585	160

字牌の間にダミーを入れて、学習を行った。学習は、±2 までの隣合った要素の組み合わせも素性としているので、ダミーは 2 つずつ入れて実験を行った。

4.4 周辺確率を用いた不完全情報推定

CRF を用いた不完全情報推定では、手牌が 13 枚になることがほとんどなく、推定した手牌の枚数はほぼ 0 枚となっていた。

これを改善するために CRF++ で得ることができる、周辺確率を用いて不完全情報推定を行う。CRF++ で得ることができる、周辺確率を用いて不完全情報推定を行う。周辺確率とは、同時分布 $P(x, y)$ があつたときに、それを y について足しあわせ得られる $P(x)$ であり、ラベルの付く確率である。本実験におけるラベルは枚数を表すので、周辺確率はある牌を n 枚持っている確率に相当する。ある局面における周辺確率を 5 に示す。

表 5 から、推定する相手が東を 0 枚持っている周辺確率は 0.922745%、1 枚持っている周辺確率は 0.053877%、2 枚持っている周辺確率は 0.020137%、3 枚持っている周辺確率は 0.003121%、4 枚持っている周辺確率は 0.000120% とする。表 5 から、推定する相手が東を 0 枚持っている周辺確率は 0.922%、1 枚持っている周辺確率は 0.0538%、2 枚持っている周辺確率は 0.0201%、3 枚持っている周辺確率は 0.00312%、4 枚持っている周辺確率は 0.000120% となる。

周辺確率を用いた不完全情報推定では、手牌が 13 枚になるまで、ラベル 0(0 枚持っている状態)の周辺確率が小さい牌から順にラベル 1 から 4 をルーレット選択する。

表 5 周辺確率の例

牌	ラベル 0	ラベル 1	ラベル 2	ラベル 3	ラベル 4
東	0.922	0.0538	0.0201	0.00312	0.000120
南	0.960	0.0304	0.00491	0.00452	0.000088
西	0.969	0.0187	0.00785	0.00354	0.000065
北	0.959	0.0220	0.0122	0.00575	0.000076

5. 不完全情報推定の実験

学習データは牌譜からランダムに抽出した 6039 局面に対して

- (1) ダミーなし
 - (2) ダミーあり (ダミーは牌種: dm, ラベル: 0)
 - (3) ダミーあり (ダミーは牌種: dm, ラベル: d)
- の 3 つを用意した。

ダミーのラベルにはラベル 0, ラベル d を用いた。

表 6 に周辺確率を用いた不完全情報推定のラベル一致率を示す。

図 3 に実際の手牌と周辺確率とダミーを用いて生成した手牌のシャンテン数の推移を示す。

データ名	総数	一致数	一致率 [%]
ダミーなし	360661	148994	41.3
ダミー 0	360661	150105	41.6
ダミー d	360661	150470	41.7
ランダム	360661	113584	31.5

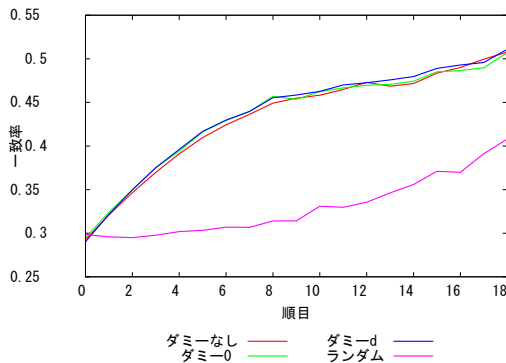


図 2 ダミーを用いた不完全情報推定の一一致率の推移

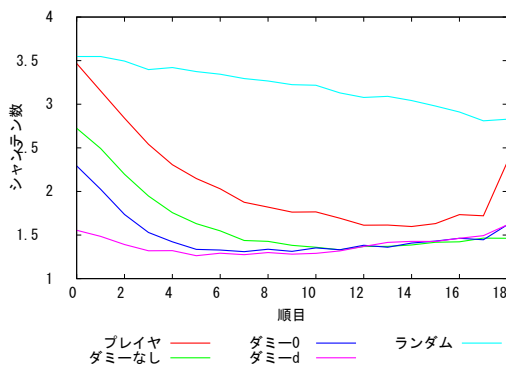


図 3 実際の手牌と推定した手牌のシャンテン数の推移

6. 考 察

提案手法の手牌の一致率は 42% となったが、ダミーを用いることによる優位性は見られなかった。順目が進むごとに一致率が増えていくことは、場に見える牌の数が多くなるからだと考えられる。

またシャンテン数の推移について、提案手法で推定した手牌のシャンテン数は実際のプレイヤーのものよりも低くなった。これは、順目の情報を特徴量として使わずに学習したため、手牌に残りやすい牌、面子の作りやすい牌が選ばれやすくなるように学習されたからだと考えられる。

7. おわりに

本論文では CRF の着想を用いて、相手の河にある枚数や場に見えている枚数から、相手の手牌の推定を行うシステムを制作した。

今回は、一致率とシャンテン数で評価した。しかし、一致率は他に基準がないため、これだけでの評価は困難である。

今後の予定として、別の評価方法での評価や、手牌の生成方法を改良することが挙げられる。

参 考 文 献

- 1) 三木理斗, 三輪誠, 近山隆: UCT 探索による不完全情報下の行動決定, 第 14 回ゲームプログラミングワークショップ, pp. 43-50, 2009.
- 2) Lafferty, J., McCallum, A. and Pereira, F.: Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data, In Proc. of 18th International Conference on Machine Learning, pp.282-289 (2001).
- 3) 高村大也 著, 奥村学 監修: 言語処理のための機械学習入門, コロナ社, 2010.
- 4) CRF++, <http://crfpp.sourceforge.net/>
- 5) とつげき東北: システムティック麻雀研究所, <http://www.interq.or.jp/snake/totugeki/>