

進捗報告

1 今週やったこと

- 実験結果

2 評価指標

今回の実験の評価指標はmapを参考にした。すべての手牌に対してこの一枚を捨てる時の有効牌の数を計算して、シャンテン数が一番少ないのだけを、ランク付ける。実験結果のランキングを実際のランキングと比較して、評価指標を用いて評価する。また、今回は二つの評価指標を使用する。一つ目は一番多い選択だけが評価対象、二つ目はTOP3が評価対象。具体的な計算方法を示す。例えば、実際の1位の牌は今回のモデルの計算で2位になった場合は評価点数は1/2になる。だが、二つ目の評価指標に対して、2位の牌は1位に計算された場合は減点なくて評価点数は1になる。

3 実験結果

表1に結果を示す。「残り牌」はプレイヤーが見えない牌で計算した結果で今回のBaselineになる。結果1は0-4の数字五つしかないが、結果2はこの五つの数字かける各自の確率で計算した。全体の結果を見ると、今回結果はBaselineを上回ったが、僅かの差しかない。また、テストデータは3種類に分かれた。一つ目は選択は一つしかない場合、この場合は計算する必要がない。二つ目は選択は二つの場合、この場合はどういう評価手法を使用するのかはまだ考えている。三つ目は選択が三つ以上選択がある場合、表2に選択が三つ以上あるの結果を示す。また、この三種類は今回のデータで5758:4302:32632になり、合わせて42692個のテストデータになる。

表 1: 全体の結果

	TOP1	TOP3
残り牌	0.8608	0.8709
結果 1	0.8656	0.8770
結果 2	0.8043	0.8256

表 2: 選択三つ以上あるの結果

	TOP1	TOP3
残り牌	0.8227	0.8319
結果 1	0.8286	0.8398
結果 2	0.7519	0.7729

4 気になる点

- モデルの予測データは残り牌の数より多いデータはある。
- 字牌はほぼ 0 と予測された。
- 評価手法はランキングしか見てないが、有効牌の数もかなり重要だと考える。例えば、結果とラベルの差や、1 位と 2 位の差など。

5 今後の予定