

# 由不考虑多玩家因素的监督学习实现四人麻将程序

水上直紀<sup>1,a\*</sup>, 中張遼太郎<sup>1,b</sup>, 浦晃<sup>†1,1,c</sup>, 三輪誠<sup>†2,2,d</sup>, 鶴岡慶雅<sup>1,e</sup>, 近山隆<sup>1,f</sup>

## 摘要

本论文阐述了由使用牌谱的监督学习生成麻将程序的方法。首先，使用高级别玩家的牌谱的感知器学习生成单人麻将程序，再对其扩展以适应四人麻将。扩展是在单人麻将程序中加入了‘弃和’和‘鸣牌’能力的监督学习成果。联机麻将网站‘天凤’上对生成的程序进行评价，其R值为1651，显著高于一般人类玩家。

## 关键词

多玩家游戏 麻将 监督学习

<sup>1</sup> 東京大学 The University of Tokyo

<sup>2</sup> マンチェスター大学 The University of Manchester

<sup>†1</sup> 現在，富士通研究所

<sup>†2</sup> 現在，豊田工業大学

<sup>a</sup> mizukami@logos.t.u-tokyo.ac.jp

<sup>b</sup> nakahari@logos.t.u-tokyo.ac.jp

<sup>c</sup> ura@logos.t.u-tokyo.ac.jp

<sup>d</sup> makoto-miwa@toyota-ti.ac.jp

<sup>e</sup> tsuruoka@logos.t.u-tokyo.ac.jp

<sup>f</sup> chikayama@logos.t.u-tokyo.ac.jp

\*译者：零之审判 北京大学数学科学学院 学士毕业

## 1. 前言

近年二人零和的完全信息博弈游戏，如国际象棋、黑白棋、将棋，已经实现了和人类顶尖玩家同等或更强实力的电脑程序[1][2][3]。另一方面，对于扑克和麻雀为代表的完全信息博弈多人游戏，实现不逊于人类顶尖玩家实力的电脑程序例子还非常之少[4][5][6][7]。

作为不完全信息博弈游戏的一个代表，德州扑克（扑克的一种）中提出了依靠自我对战计算的近似纳什均衡策略，能生成2人游戏下一个比冠军人类玩家更强的电脑程序[8]。另外，2人游戏得到的纳什均衡策略扩展到3人游戏的方法，以及通过简化玩家的行为假想少量玩家的少量对战的方法[5]也被提出。

本论文研究对象‘麻将’是4人零和不完全信息博弈游戏。至于生成电脑麻将程序的方法，目前提出的有完全学习4人麻将牌谱的程序[6]和使用蒙特卡罗树搜索模拟所有玩家打牌的程序，但其都是大大弱于一般人类玩家的实力水平。

本论文提出的麻将程序，图1展示了其生成流程。首先，我们通过高级别玩家的牌谱进行监督学习，生成将麻将的多玩家因素排除掉的单人麻将程序。然后，为了消除单人麻将程序在四人对战下的弱点，加入了‘弃和’和‘鸣牌’的能力。最后我们在联机麻将网站上与人类玩家对战，对总程序进行评价。

本论文的框架如下。首先是第二章的麻将规则和术语，第三章阐述相关研究。程序的提出部分，

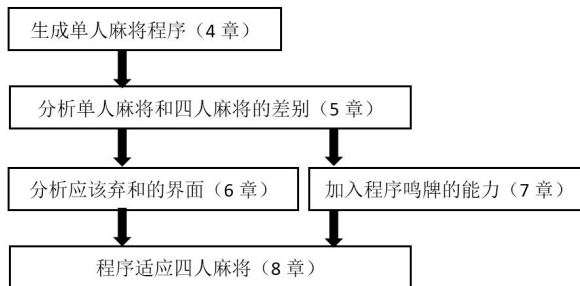


图 1. 提出的手段

第四章生成单人麻将程序、第五章分析单人麻将和四人麻将的差别、第六章分析应该弃和的界面、第七章加入程序鸣牌的能力、第八章说明程序适应四人麻将的方法事项。最后是第九章的总结陈述。

## 2. 麻将规则和术语

麻将是由4位玩家互相争夺点数的游戏。

每局开始时各个玩家手中的13张牌称为配牌。之后各玩家按顺序摸上1张称为自摸牌的牌、并切出一张牌。手牌的13张，加上自摸牌或他家切出的牌组成的14张有役（特别的牌型）时，玩家和了，并得到与役对应的点数。和自摸牌称为自摸，和他家切牌称为荣和。而被荣和则称为放铳，和了点数全部由放铳玩家支付。一次麻将的对局进行特定数量的轮回（常见的是八轮），每轮即从配牌到和牌，玩家们竞争对局结束时的点数高低。

14张牌的组合符合和牌条件，需要四个特别组合的三张牌作为面子和相同的两张牌作为对子。玩

家可以利用他家切出的牌组合面子，被称作**鸣牌**或**副露**。**搭子**指加入一张牌可以组合成面子，即可以通过鸣牌生成面子。

麻将由136张牌组成，34种不同的牌、每种4张。其中108张为**数牌**，数牌中包括三套1到9，分别为**万字**、**筒子**和**索子**，三种没有优劣之分。**字牌**由风牌（东、南、西、北）和三元牌（白、发和中）组成。

另外说明以下的常见麻将术语。

**立直** 没鸣过牌，且手牌只差一张就能和了时宣告的行为。宣告立直后不能改变手牌，即摸到不能自摸的牌时必须直接切出。

**七对子** 一种役，指7个对子组合的牌型。

**国士无双** 一种役，各组数牌的1、9和字牌全部有1张以上

**听牌** 加入一张牌即可和了的状态

**向听数** 到听牌状态至少需要加入的牌数

**役牌** 相同三个即可作为役的牌

**宝牌** 每局开始前随机确定的牌，和了时有该牌即可增加得点。

**牌山** 自摸牌堆在一起

**弃和** 放弃追求和牌，避免放銃

**兜牌** 不选择追求和牌的最优切牌，同时避免放銃和试图听牌的策略

**现物** 对手切出的所有牌，根据规则现物是不会放銃的。

**筋** 与现物的数字相差3的靠近1、9的数牌

**壁** 之间的数牌已经没了

**染手** 役的总称，指数牌全是一套的牌型

**断么** 一种役，没有数牌的1、9和字牌的牌型

**杠** 一种鸣牌，将4张相同牌组合的行为，此时场上增加1种宝牌

**吃** 一种鸣牌，生成相连3张数牌组合的面子

**碰** 一种鸣牌，生成相同3张牌组合的面子

**暗刻** 不鸣牌得到的相同3张牌组合的面子

### 3. 相关研究

作为纳什均衡策略适用于多人博弈游戏的研究，已经有益于扑克[4][5]。二人扑克下使用近似纳什均衡策略已经战胜了世界冠军，但玩家数量增加时状态数指数上升，因此纳什均衡策略的计算变得困难。于是2人游戏得到的纳什均衡策略扩展到3人游戏的方法[4]，以及通过简化玩家的行为假想少量玩家的少量对战的方法[5]也被提出。然而，存在的问题是

无法剥削偏离纳什均衡的弱者，以及太过简化而不能适应于现实游戏。

麻将方面，北川他们提出了使用3层神经网络的麻将程序[6]。在牌谱中何切的局面和鸣牌的局面寻找自己手牌的正确何切，使用3层神经网络学习出应对下次局面的选择。与人的对战行为中，因为没有针对的学习牌的组合，有和牌技术与防守技术低下的问题。

三木他们提出了使用蒙特卡洛搜索的麻将程序[7]。这种手段为，由于麻将游戏树的搜索分支数特别的大而很难正确前进，因此使用将对手的手牌和切牌随机化模拟的蒙特卡洛树搜索[9]。尽管几乎没有使用麻将的知识，该程序仍然比仅仅试图减小向听数的简单程序表现更出色。但是该程序有与人类玩家的打牌迥异的行为，比如很难有役时鸣牌，因为模拟中对手玩家几乎不会和牌的。

上述的所有研究中，生成的电脑玩家在東風莊[10]中只有倒数13%成绩的实力，没有达到人类玩家平均水平。

以下举例在游戏中适用机器学习技术的代表研究。黑白棋中以盘中的形状作为特征量，从对局结束的胜负学习其权重[1]。还有在西洋双陆棋中使用TD学习[11]，这是强化学习的一种，为没有结果的局面下亲局面的评价价值尽量接近子局面的评价价值的学习方法。然而，麻将并不是最优解就能和牌，难以直接使用这些方法。

## 4. 生成单人麻将程序

本章说明单人麻将程序，这是扩展到四人麻将的基础。

### 4.1 单人麻将

四人麻将中战绩优秀的玩家，统计得知其平均和了点低而和率[12]。强者们平均和了点低的最大原因是其不会强要手役的打牌方式。麻将点数的特征是，一定点数以上的手役不会因为难易度增加点数，因此多用点数较低的役多和牌的效率更好。而且，通过自己的和牌减少对手和牌机会，这种想法来看和率是重要的。

本文的研究中，为了生成相比平均和了点更重视和率的程序，对单人麻将程序的评价只用是否和了来判断。以下为单人麻将的规则：

- 只能自摸和牌
- 因为没有对手，对手切牌也没有

- 没有鸣牌和立直

## 4.2 基于平均感知器算法的单人麻将程序

本节陈述单人麻将程序的生成方法。本研究采用平均感知器算法[13]的监督学习生成单人麻将程序。感知器算法对于有相关性的大量特征量的学习，有着相对较小的计算耗时的优势。

以下 $\mathbf{x}$ 为局面（麻将游戏界面）的特征向量， $\mathbf{w}$ 为权重向量， $n$ 为向量长。 $x_i$ 为特征向量的第 $i$ 元， $\omega_i$ 为权重向量的第 $i$ 元，任给的局面下得分由式（1）计算。

$$f(x, w) = \sum_{i=1}^n x_i \omega_i \quad (1)$$

14张牌中去除一张后的13张，对其提取特征量，由此计算各种切牌下的得分。再将牌谱中实际切牌 $t^*$ 与得分最高的切牌 $\hat{t}$ 比较，两者不同时，使用 $t^*$ 下的特征向量 $\mathbf{x}_{t^*}$ 和 $\hat{t}$ 下的特征向量 $\mathbf{x}_{\hat{t}}$ 来更新特征向量 $\mathbf{w}'$ 由式（2）计算。

$$\mathbf{w}' = \mathbf{w} + x_{t^*} - x_{\hat{t}} \quad (2)$$

对牌谱的所有界面重复进行更新，每次得到的权重向量的平均值作为最终的权重向量。

使用的特征量是参考北川他们论文[6]设计得到（表1），特征向量的长度为37514。感知器学习中训练数据遍历更新3次。

## 4.3 学习的牌谱

训练数据使用的是在线麻将网站天凤[14]的凤凰桌牌谱<sup>1</sup>。能在凤凰桌对局需要在全体玩家中排位0.1%的程度，可以认为牌谱质量很高。然而牌谱中含有许多四人麻将特有的局面和切牌，直接作为训练数据是不恰当的。因此我们确认每局先制立直的玩家，只将其的切牌作为训练数据，而且局面为配牌开始直到其立直的游戏界面。最终得到的训练数据约有170万个界面。

## 4.4 单人麻将的实力

为了验证成功的训练，我们评价生成的单人麻将程序。为了去除其他玩家的影响，测验是程序单人打牌。而且为了尽量减小运气的因素，用同样的牌山来比较多个单人麻将程序的实力。测验中使用不同的100个牌山，评价了各个玩家的打牌结果。玩家包括‘高级玩家（人类）’和‘一般玩家（人类）’

各一人，以及前人研究[7]用到的Plain UCT。一般玩家是天凤[14]中排位大约50%的玩家，高级玩家是本文的第一作者，在天凤中排位前0.1%。

表2展示了结果。单人麻将程序的胜率比一般玩家要高，考虑到前人研究远远没达到一般玩家的水平，可以说单人麻将程序因为针对性地学习了构造的存在特征量，得到了实力的飞跃提升。

将单人麻将程序与小松他们设计的程序[15]进行比较。然而小松他们使用了宝牌，以结果的点数来进行评价，因此和本研究的设定有些不同。小松他们研究提出的方法是在手牌中加入一定牌后，选择最大和了点的全部组合，在手牌中全部组合不需要的牌上加上和了点，以此来选择切出的手牌。

使用1万局不同的牌山对两个程序比较和率和平均和了点，表3展示了结果。虽然平均和了点劣于小松们的研究，但和率更高，显示我们研究中对单人麻将程序使用牌谱训练是成功的。

## 5. 分析单人麻将和四人麻将的差别

前章展示了对于单人麻将，通过使用牌谱的监督训练生成了与一般人类玩家实力相近的程序。本章为了分析单人麻将和四人麻将的差别，在四人麻将的牌谱中调查程序的切牌与谱中切牌的不同。

前章生成的单人麻将程序没有‘鸣牌’的功能，而实际牌谱中和了80万个局面<sup>2</sup>中，没有鸣牌即和了的比例为52%，鸣牌后和牌的比例为48%。从该结果来说可以认为不能鸣牌的影响是非常大的，但本章只分析自摸牌的局面下单人麻将与四人麻将的不同。

具体分析如下。首先在实际牌谱的界面下由单人麻将程序选择前3位想切出的牌，收集牌谱中的实际的切牌不在这3位中的界面，人工标记‘弃和’和‘兜牌’，统计其个数，从统计的结果分析单人麻将和四人麻将的差别。

首先评价单人麻将程序在四人麻将里牌的选择准确度。这里程序指定的切牌候补、即想切出的牌前 $n$ 位中包含牌谱里实际切牌的比例定义为Rank  $n$ ，作为比较也对高级玩家进行评价。

测试集是从天凤[14]公开的凤凰桌牌谱里随机抽出的27局中的1342个界面。为了匹配单人麻将，1342个界面中没有玩家鸣牌或鸣牌之后的界面。单人麻将中是没有对手的切牌和副露牌的，但测试集

<sup>1</sup>2009年2月20日至2010年1月31日的对局

<sup>2</sup>2010年1月1日到2010年12月31日的对局，获得了天凤开发者角田的许可

表 1. 单人麻将程序的特征类

特征类	特征个数
面子手、七对子、国士无双向听数	$15+7+14=36$
3色（万、筒、索）中不使用n色（ $n=1\sim 3$ ）的n和向听数	$3\times 16=48$
色的个数（ $1\sim 4$ ）及各色向听数的组合	$12+78+364+1365=1819$
各色牌中最多的数量	15
各色牌中最多的数量+字牌数	15
2到8的数牌量	15
面子和不含有幺九的连续2张牌的搭子和面子候补的组合	$6+12+72=90$
各字牌的枚数和役牌的枚数和是否是宝牌的组合	$5\times 6\times 2=60$
数牌的数字-5的绝对值和枚数和是否是宝牌的组合	$5\times 5\times 2=50$
各色牌中是否有 $1\sim 9$ 的组合	512
连续的n个数牌中各牌数量的组合（ $n=2\sim 6$ ）	$100+500+1860+8634+23760=34854$

表 2. 单人麻将打牌和率

玩家	和率(%)	平均和了巡目
高级玩家	17	13.8
单人麻将程序	13	12.7
一般玩家	12	12.8
Plain UCT	2	15.0

表 3. 1万局的和率和平均点数

玩家	和率(%)	平均和了点
单人麻将程序	20.4	733
小松们的研究[15]	14.0	850

中玩家利用了这些牌计算剩下枚数。高级玩家（人类）的选择按照四人麻将的标准，单人麻将程序的选择按照训练的单人麻将的标准，表4展示了评价结果。

表 4. 和牌谱的一致率

玩家	Rank 1	Rank 2	Rank 3
高级玩家（人类）	0.62	0.85	0.93
单人麻将程序	0.53	0.77	0.85

高级玩家的Rank 1的正解率为62%，单人麻将程序为53%，测试集不同时报告的精度为56%。

然后是调查单人麻将程序的Rank 3里不含有正解（牌谱的切牌）的界面，即剩下的15%的193个界面，人工对其标记标签进行分类。图2展示了结果。这些界面里大约4成为应当弃和的局面，还有很多做‘役牌’、‘七对子’、‘染手’、‘断幺’、‘手役的界

面，还有一些是单人麻将程序本身的恶手（错切）。

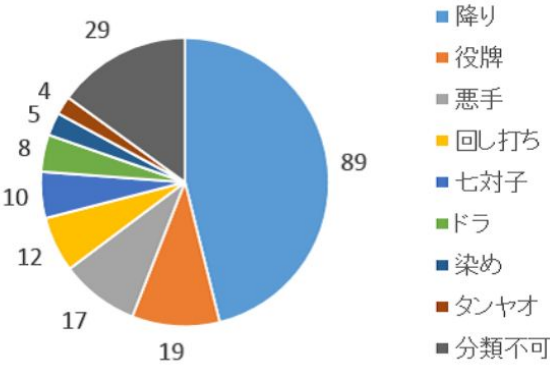


图 2. 没命中正解的界面分类

对分类的标签进行说明：手役指的是单人麻将程序没有注重该手役的界面；‘宝牌’指的是切了没必要切的宝牌的界面；‘分类不可’指的是候补的3个切牌一样、没有选择正解的界面或牌谱本身恶手的界面。

根据以上的分析，可以知道弃和是单人麻将扩展到四人麻将时最重要的问题，如果能判断弃和，预测与牌谱的一致率能上升15%中的4成、即6%。

## 6. 识别应该弃和的界面

根据前章的分析，正确识别弃和的界面能提升牌谱一致率，增强实力。因此这章首先说明为了识别弃和的界面而对其标记的标签，然后说明识别的提出手段。

### 6.1 训练集生成

监督学习下识别弃和的界面需要训练集，然而



牌谱中没有明示‘弃和’。因此人工标记标签，此时是否是弃和的判断很重要。

图3展示了具体例子。图中界面玩家（界面下方）切出了3万，也就是说从面子里切牌从而放弃了手牌和牌可能性的选择。在标记标签时，只要一次弃和就把该局之后全当做弃和。



图 3. 弃和的界面

根据以上的定义，第一作者（A）从牌谱中的509局里手工标记了5716个界面是否是弃和的标签，即观察牌谱里的切牌来标记标签。

为了验证标记标签的妥当性，第二作者（B）独立地标记了5716个界面中的一部分、1053个界面。第二作者在麻将网站‘天凤’[14]中排位前10%。

表5展示了两人的标记标签的一致率。由于判断受玩家主观想法影响很大，因此是困难的问题。但一致率达到了了一定的值，排除随机性的一致率指标kappa系数[16]计算结果为0.75，可以认为是否弃和的界面判断大体上一致。观察不一致的48个界面，发现兜牌的界面有47个，剩下的一个是牌谱明显的恶手。

表 5. 标签标记的一致率

A/B	弃和	非弃和
弃和	83	23
非弃和	25	922

图4展示了兜牌的具体例子。不管其他玩家的话最佳解应当是8索，但界面里对付右边玩家的立直切出了现物1索。

而且就算能判断其非最佳解，但很难判断是否是兜牌的界面也存在。图5展示了两人的判断不一致的



图 4. 兜牌的界面

界面。牌谱的切牌为2万，A认为无论怎么看都是拆了面子因此判断为弃和。然而还有其他看法，2万和6万是右边立直玩家的现物，4万也是较安全的筋牌，把这些牌切出的时候可能剩下的形状摸成面子或立直玩家切出了看起来危险的牌，可以说玩家还没有放弃增加点数，B因此判断为兜牌。

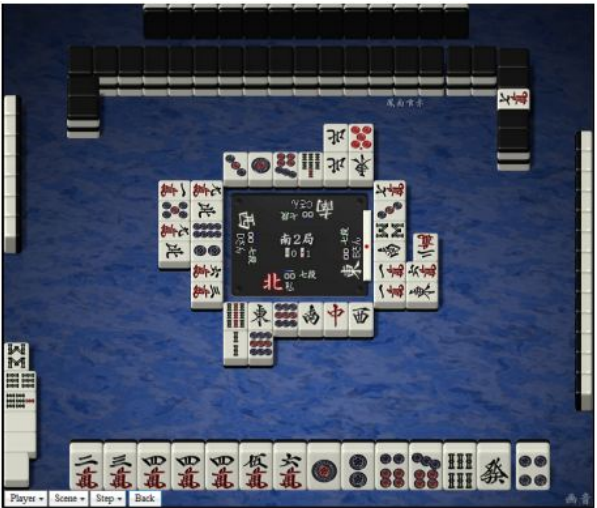


图 5. 难以判断弃和或兜牌的界面

像这样的切牌，是否试图和了或流局听牌的判断受玩家主观想法影响很大。特别是七对子和牌可能性看起来多大对玩家来说是很困难的。

另外，因为判断弃和了的场合根据定义之后全部判断为弃和，所以对于弃和和兜牌的判断不一致的场合，当之后的自摸牌是安牌时，一方判断弃和，另一方判断不弃和。因此一旦判断不一致，之后的判断就有很大的分歧，导致一致率低下。

6.2 识别应该弃和的界面

本节说明监督学习识别应该弃和的界面的方法。首先说明从界面里提取什么样的特征类，再评价识别准确率。

本节的问题设定和6.1节有些许不同，6.1节是观察牌谱的切牌来判断是否是弃和，然而实际中需要在切牌前判断是否弃和，根据判断来选择切牌。因此与6.1节不同，我们对界面是否弃和进行判断。

表6展示了判断是否弃和时从界面提取的特征类。因为需要考虑自身与三家的状态和场上的情况，特征向量长度为34（7+3×8+3）。

表 6. 弃和判断的特征类

	项目	数量
自身状态	向听数	1
	七对子向听数	1
	是否立直	1
	是否坐庄	1
	副露数	1
	副露的宝牌数	1
	手牌宝牌数	1
对手状态	是否立直	1
	副露数	1
	副露可见的宝牌数	1
	是否坐庄	1
	想切的牌对该玩家安全度	4
场上状况	牌山剩余量	1
	杠的数量	1
	牌河里宝牌数量	1

为了简化研究的问题，弃和的判断基于以下的模型考虑。人类玩家决定是否应该弃和时，会先选择不考虑他家的最佳切牌，然后考虑这张牌对其他玩家的是否安全来判断是否弃和。因此模型里要考虑想切的牌对他家的安全度。

安全度由启发式决定，其高低顺序如下：

- 现物
- 现1枚以上的字牌和1，9筋牌和壁
- 字牌和其他筋牌
- 无筋

6.3 识别弃和界面的预实验

为了验证6.2节特征类的妥当性，我们进行识别弃和界面的2元分类。学习弃和判断的训练集相对

较小，因此研究使用的分类器为高斯核的Support Vector Machine（SVM），它相比准确度最大化能得到更高的泛化性。

评价使用的5折交叉验证。使用封装SVM的LIBSVM[17]，根据grid search寻找Gaussian核下最适合的参数cost（c）和gamma（g），搜索的范围为两参数从 $2^{-10}$ 到 $2^{10}$ 的刻度为2倍的网格。

训练数据中弃和的界面为531，非弃和的界面为5185。表7展示了结果。表中的准确度为程序输出的预测为实际正解的比例，测试准确度为测试集中预测为实际正解的比例，两者的调和评价为F值。根据网格搜索， $c = 2^8$ ， $g = 2^{-7}$ 时弃和界面的F值最高。

表 7. 弃和界面的分类结果

	弃和界面	非弃和界面
准确度	0.76	0.97
测试准确度	0.71	0.97
F值	0.73	0.97

由结果可知对弃和预测的精度相对较高。

7. 加入程序鸣牌的能力

本章将单人麻将程序扩展为能够鸣牌，而且以牌谱为准确定鸣牌判断的准确度。

7.1 基于平均感知器算法对鸣牌局面的学习

对鸣牌的学习使用和4.1节用到的平均感知器算法。鸣牌后能切出的牌，不能用其所在已经完成的面子去鸣牌（比如123吃4后切出1，这是禁止的）。因为有这样的规则，所以与自摸牌局面不同。鸣牌和自摸牌局面的学习中的不同点是，比较的是能鸣牌的局面下不鸣牌的手牌和鸣牌后切出一张牌的手牌的得分。特征类在4.1节使用的特征类外增加了与鸣牌相关的部分。（表8）

特征向量长度为69258，感知器学习中训练数据遍历更新1次。

7.2 用来学习鸣牌的牌谱

4.1节使用的训练数据是立直玩家的界面，然而这里面不含鸣牌的局面，因此学习鸣牌的训练数据从4.1节使用的变更为能够鸣牌的。各局中先制立直的玩家宣示立直前，或者和了的玩家到和牌前的界面作为训练数据，最终得到的数据集有300万个界

表 8. 单人麻将程序鸣牌情况下增加的特征类

特征类	特征个数
副露数	5
是否可能立直	2
各色的1~9数牌数量变换到0, 1, 2以上的牌型	19472
副露里选择n (1~2) 个副露的种类组合	$136+9316=9452$
手牌里有3张以上相同牌的数量和2张以上相同牌的数量的组合	$5\times 7=35$
手牌里有3张以上相同牌的数量+碰的个数和2张以上相同牌的数量的组合, 还有是否吃牌	$5\times 7+1=36$
手牌和副露中断幺牌数量和向听数	$16+16=32$
连续的各色的3张数牌里的有无和3张里是否包含幺九和是否包含吃的组合	$512\times 2\times 2=2048$
包含吃的各色里幺九的有无	512
各色牌中最多的数量+该色和字牌的副露数 $\times 3$	16
各色牌中最多的数量+字牌个数+该色和字牌的副露数 $\times 3$	16
做特定色染手的向听数	16
役牌对子数	8
役牌刻子数	6
宝牌个数 (0, 1, 2, 3以上) 和2~8数牌中宝牌数量	$4+4=8$
副露数和向听数的组合	$5\times 15=75$
向听数和有役、有门清限定役、没役和巡目的组合	$2\times 3\times 19=114$
有役和是否是门清限定役和没有役	3
是否鸣了幺九牌和加入幺九牌时向听数减少的量 (0, 1, 2以上)	$2\times 3=6$
手牌里不使用幺九牌的向听数增加的量 (0, 1, 2以上)	3
加上所有副露的风牌后其数量 (0, 1, 2, 3以上) 的组合	$4\times 4\times 4\times 4=256$
加上所有副露的三元牌后其数量 (0, 1, 2, 3以上) 的组合	$4\times 4\times 4=64$
副露数和向听数和加入已经切出的牌是否减少向听数的组合	$5\times 4\times 2=40$

面。测试集则限定为能够鸣牌的约3万个界面, 并准备3000个界面作为启发集。将训练集分割成三十部分, 每使用一部分观测其对启发集的正解率。

7.3 鸣牌的正解率

为了验证针对性的进行了对鸣牌的学习, 调查程序对开发集的学习曲线和测试集的正解率, 图6和表9展示了结果。

表 9. 是否鸣牌的正解率

牌谱的选择	鸣	不鸣
牌谱数量	6230	22164
完全一致的数量	3755	19392
只有鸣是正解	1435	N\A
正解率	84.2	90.7

为了不改变是否鸣牌的正解率, 训练集的的遍历次数为1回已经足够了。这个鸣牌分类器对鸣牌局

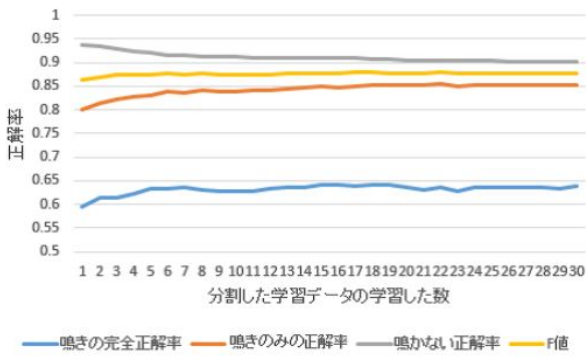


图 6. 学习曲线

面有相对高的正解率, 虽然数据集不同不能简单地比较, 但相比北川们的研究[6]里鸣牌局面下21%的正解率已经有了很大提升。鸣牌局面中我们程序正确率的大幅提升的关键原因是学习的界面并非全体界面, 而限定为正在进攻的局面。



## 8. 单人麻将程序适应四人麻将的方法事项

到前章为止，我们已经训练了对弃和局面的识别和鸣牌的机能。这一章将其与单人麻将程序组合，适应实际的四人麻将，并对其实力进行评价。

对弃和局面的识别与单人麻将程序的组合，首先说明游戏里轮流的行为。如果能荣和对手切出的牌，全部荣和。图7展示了轮到自已摸牌时的流程图。先判断能否和牌，能和时全部自摸和了，不能和时由单人麻将程序选择一张牌，以切出该牌为基准得到6.2节说明的特征量，以此判断是否弃和。如果不弃和，则按照规则切出选择的牌。如果弃和则选择弃和必要的切牌，因为现在的目的是是否弃和，所以没有考虑兜牌的选择。弃和的切牌由以下准则决定：

- 对各位玩家选择6.2节的安全度下顺位高的牌
- 多玩家立直の場合，选择对其均安全度高的牌
- 没有玩家立直の場合，关注鸣牌数多的玩家

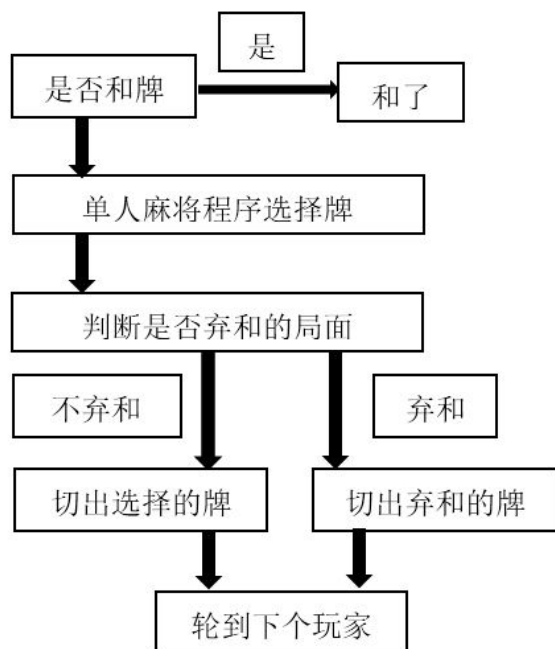


图 7. 轮到自已时的流程

图8展示了轮到对手摸牌时的流程图。首先能荣和对手的切牌时和了，然后能鸣该牌时，如果自身没有选择过弃和则进行鸣牌判断。此时判断为鸣牌时，决定鸣牌和之后切出的牌。

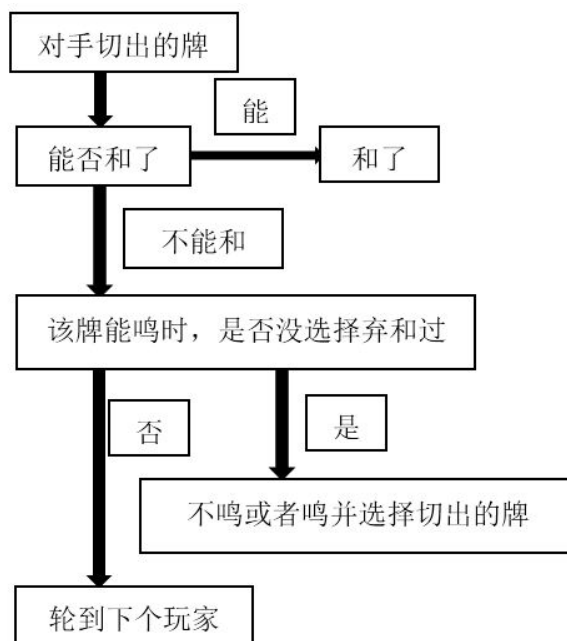


图 8. 轮到对手时的流程

### 8.1 评价设定

为了测定这个程序的实力，我们在联机麻将对战网站天凤[14]上对战。规则为东风战、有赤宝牌、摸牌切牌限时5+10秒、鸣牌选择的限时同样、鸣牌考虑时对局停顿。天凤玩家根据其战绩在不同的桌对战，共有四种对战桌，级别从高到低分别为凤凰、特上、上级和一般。虽然成绩足够进入上级桌，我们限定在一般桌进行对战。为了程序在网页上对战，我们自制了输入输出的接口。我们基于平均顺位（R值）进行评价。R值的变化与平均顺位负相关，其计算如下：

$$R = R + (50 - Rank \times 20 + \frac{AveR - R}{40}) \times 0.2, \quad (3)$$

其中Rank为该局顺位，AveR为该局全部玩家的平均R值，R初值为1500。平均顺位进步0.1约能导致R值陡增100。

### 8.2 评价设定

第4章生成的单人麻将程序作为‘基线’，与‘基线’加入弃和、再加入鸣牌的程序共3个进行比较。表10展示了结果。根据Welch’s t检验，1%的显著水平下认为任两个程序的平均顺位有差异。

表11展示了和率和放铳率的结果。和率同4.1节的说明一致，越强的程序倾向于越高。放铳指切出了对手的铳牌，因为放铳时要支付对手的全部点数，所以原则上放铳率越低越好。1局中的收支受和率和



表 10. 顺位分布

程序	1位率	2位率	3位率	4位率	平均顺位	对战数	R值
基线+弃和+鸣牌	0.236	0.321	0.234	0.209	2.41	402	1651
基线+弃和	0.237	0.240	0.259	0.264	2.54	834	1507
基线	0.181	0.216	0.252	0.251	2.77	504	1262

放铙率影响，而1局多次收支结束后确定最终顺位，因此和率和放铙率与顺位分布关系很大。

表 11. 和率●放铙率

程序	和率	放铙率
基线+弃和+鸣牌	0.275	0.149
基线+弃和	0.181	0.144
基线	0.188	0.190

‘基线+弃和’相比‘基线’程序，和率略低但铙率大幅下降。因为‘基线+弃和’程序在和牌困难的场合下察觉他家的做牌时拆了手牌、为了不放铙而弃和，所以预期上和率降低的同时放铙率也会降低。结果正符合我们预期。

‘基线+弃和+鸣牌’相比‘基线+弃和’程序，铙率略高但和率大幅上升。鸣牌后手牌数下降，因此铙率上升是预料到了的。但鸣牌后和率也更好了，结果符合我们的预期。

‘基线+弃和+鸣牌’程序得到的R值为1651，可以说强于1500点的平均玩家。虽然程序不懂手役的知识，对手的点数状况也完全不考虑，但还是达到了这样的实力，可以认为这说明了单人麻将的必要性和弃和、鸣牌的重要性。

8.3 考察

通过对‘基线+弃和+鸣牌’程序打牌的观察，其察觉他家的做牌而弃和的局面判断大体上是正确的。这些局面他家很快和牌而自家和牌困难且剩余巡目不多。

问题是第一作者判断应该弃和的局面下，‘基线+弃和+鸣牌’程序很多都没有弃和。这些场合往往是手牌向听数小，但进张枚数少因此手牌不好。其原因是对手牌好坏的判断只基于向听数和宝牌持有的枚数，这些作为判断手牌好坏的基准虽然很重要，但并不是判断的全部。问题的原因出在与弃和相关的学习样本太少，加入手牌的信息后预实验的结果也不好从而没有加入向听数以外的判断基准。

进一步提高弃和选择的准确度，分析对手的

铙牌时必须的。现在的手段只是将牌的安全度分成4级，以后考虑从对手的切牌中预测各个牌的安全度。比如说从对手的副露分析其手役、了解到安全牌的技术，染手之类的手役就相对容易分析。另外，牌的安全度和规则上所述有很大的不同。即使面对立直高级别玩家一般也在无筋的牌中分析其安全度，为了从中判断是否弃和，考虑的方法比规则要复杂得多。

单人麻将程序虽然比平均人类玩家的和率高，但也存在恶手。例如对子的价值，对子可以增加2枚进张，而且和牌时对子也是必需的，故对其价值的判断很困难。因此程序不怎么选择拆对子，结果是其对人战七对子的和率为3.3%，约为平均值1.7%的两倍。单人麻将程序还没有达到高级别玩家的水平，因此特征类的改良还有必要。

鸣牌的局面判断大体上是正确的。问题是训练样本中多是最终和了的局面，导致手牌难以和牌时也进行了错误的鸣牌。这可以通过在训练样本中增加没有和牌的局面来解决。

9. 总结

本研究提出的手段以下，考虑排除掉多玩家因素的单人麻将，分析其与四人麻将的差别，对单人麻将程序进行扩展以消除差别。研究生成的程序在天凤[14]对战402局，得到的R值为1651点，确认其强于平均人类玩家。考虑到还没有报告过比平均人类玩家水平要强的程序，这是一个巨大的进步。今后的课题是通过对单人麻将程序的能力改进，来改善其放铙率等。

对于弃和的学习，虽不完全但取得了一定的成果。以后的课题是兜牌的能力，现在的程序还只能把兜牌的局面看做为没必要弃和。学习了兜牌后，可以同时识别对手的进攻和进展自己的手牌，由此不仅避免切出对方的铙牌还能追求自己和牌，预期能够改善放铙率并提升战绩。

对于鸣牌的学习，和弃和一样取得了成果。以后要加入点数状况的考虑，根据点数的状况，鸣牌

后和出小牌可能是恶手。我们认为在单人麻将的特征类中加入点数状况可以避免这些恶手发生。

## 参考文献

- [1] M.Buro. Logistello: A strong learning othello program. *19th Annual Conference Gesellschaft für Klassifikation eV*, pages 1–3, 1995.
- [2] M.Campbell, A.J.Hoane Jr, and F.-h.Hsu. Deep blue. *Artificial Intelligence*, 134(1):57–83, 2002.
- [3] 田中哲朗、金子知適. 4 大規模クラスタシステムでの実行: Gps 将棋の試み (<ミニ特集> コンピュータ将棋の不遜な挑戦). *情報処理*, 51(8):1008–1015, 2010.
- [4] N. A.Risk and D.Szafron. Using counterfactual regret minimization to create competitive multiplayer poker agents. *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, 1:159–166, 2010.
- [5] 古居敬大、浦 晃、三輪 誠、鶴岡慶雅、近山 隆. 相手の抽象化による多人数ポーカーの戦略の決定. *Proceedings of the 17th Game Programming Workshop*, 2013.
- [6] 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. *Proceedings of the 12th Game Programming Workshop*, 2007.
- [7] 三木理斗、近山 隆. 多人数不完全情報ゲームにおける最適行動決定に関する研究. Master’s thesis, 東京大学, 2010.
- [8] M.Bowling, N.A.Risk, N.Bard, D.Billings, N.Burch, J.Davidson, J.Hawkin, R.Holte, M.Johanson, and M.Kan et al. A demonstration of the polaris poker system. *Proceedings of 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 1391–1392, 2009.
- [9] L.Kocsis and C.Szepesvári. Bandit based monte-carlo planning. *Proceedings of the 17th European Conference on Machine Learning*, pages 282–293, 2006.
- [10] mjman. 東風荘. <http://mj.giganet.net/>.
- [11] R. S. Sutton. Learning to predict by the methods of temporal differences. *Machine learning*, 3(1):9–44, 1988.
- [12] とつげき東北、福地 誠. おしえて!科学する麻雀. 講談社現代新書, 2009.
- [13] M. Collins. Discriminative training methods for hiddenmarkov models: Theory and experiments with perceptron algorithms. *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, 10, 2002.
- [14] 角田真吾. 天鳳. <http://tenhou.net/>, 2014.
- [15] 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, [ゲーム情報学]*, 2012(8):1–8, 2012.
- [16] S. C. Series. Measurement of observer agreement. *Radiology*, 228:303–308, 2003.
- [17] C.-C. Chang and C.-J. Lin. Libsvm. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, 2001.