

麻将AI ‘爆打’ 论文的译注

零之审判*

摘要

我翻译了麻将AI ‘爆打’ 相关的两篇文章，其分别发表在情報処理学会論文誌、CIG2015。这两篇文章是作者开发AI时最关键的阶段性成果，其背景和概要介绍如下。

* QQ 1931521908

1. 前言

麻将AI ‘爆打’ 是近年颇受关注的电脑麻将玩家，经过多次的改进，现在在天凤网站达到了特上卓安定近七段的水平。AI的第一作者水上直纪已从东京大学博士毕业，他也是天凤的凤凰玩家，现在已经入职以开发将棋AI出名的HEROZ公司。

水上直纪的主页为<http://www.logos.t.u-tokyo.ac.jp/~mizukami/>，上面列出了博士期间开发爆打时发表的所有论文，里面关键的阶段性成果按时间先后顺序为：[1]由不考虑多玩家因素的监督学习实现四人麻将程序(jousho2014)，[2]根据蒙特卡洛模拟和对手模型生成一个麻将玩家(cig2015)，[3]以推定最终顺位的期望为基础构建电脑麻将玩家(GPW2015)。第一篇为‘情報処理学会論文誌’特选论文，第二篇入选了IEEE会议，第三篇得到了优秀论文赏，我对前两篇进行了翻译，国内的同好们阅读时请按照顺序。第三篇主要是对第二篇论文中的蒙特卡洛模拟的惩罚奖励值进行了改进，没有新的切牌选择流程，所以不具体翻译了，其内容概要在后面介绍。三个阶段成果下AI在天凤的R值分别为1651、1718、1844，在此之后爆打的作者又进行了一些小改进，便达到了如今的水平。

2. 论文概要

这一部分是译者对三篇论文的概括，能够简单地了解清楚AI爆打（2015年以前的版本）在麻将对战中计算的是什么。

2.1 单人麻将打牌

单人麻将指的是不考虑对手，纯粹根据手牌选择最快可能和牌的切牌。单人麻将程序的构建在jousho2014论文的第四章，它的学习样本是天凤凤凰桌的牌谱中先制立直玩家在立直前的切牌。

2.2 鸣牌

在单人麻将程序中加入是否鸣牌的判断是在jousho2014论文的第七章，它的判断准确率（与凤凰玩家的一致率）较高。

2.3 对手模型

程序根据场上所见的数万个特征可以预测对手玩家听牌的概率，打点的高低，每张牌是其铳牌的可能性。其学习的方法在cig2015论文的第四至第六章，前两者的预测准确度与水上直纪相当。

2.4 最终顺位

程序根据某一局结束时的各家点数预测自家在东风或半庄战结束时在各顺位的概率。预测基于的特征值主要是各家的顺位、点差、剩余的局数。预测模型在GPW2015的第四章。

2.5 弃和失败率

程序会根据手牌和对手模型中对牌安全度的预测计算弃和成功率。预测函数使用的变量为对手模型里听牌概率、各张牌的放铳可能性、玩家已经看见的牌、目前的手牌，代码在GPW2015的第五章。

2.6 总流程

总流程的具体陈述在cig2015的第七章和GPW2015的第五章。爆打首先基于单人麻将程序选择一张牌，根据对手模型计算切出该牌损失的量。损失量较小，而且没有玩家立直、牌山剩余枚数较多时直接切出选择的牌。当单人麻将程序判断为鸣牌并切出一张牌时，同样根据这张牌的损失量和其他判断是否采用单人麻将程序的选择。

对于剩下的情形（这个情形比例更高），爆打会计算切出每张牌后在东风或半庄战结束时的平均顺位期望，并切出平均顺位期望最小的牌。切出某张牌后平均顺位期望的计算基于蒙特卡洛模拟，蒙特卡洛树的根节点为切出该牌的界面，蒙特卡洛树

的子节点中对手按照程序先验知识里的概率分布听牌或弃和，而自家使用单人麻将程序并且切牌损失量较大时弃和，需要注意的是自家弃和后仍然会根据计算的弃和失败率作为概率分布放铳。在某家和牌或流局后蒙特卡洛树搜索停止，根据点数变化计算最终顺位期望，作为该蒙特卡洛树的惩罚奖励值，

模拟中每个树的惩罚奖励值的平均即为切出该牌后在东风或半庄战结束时的平均顺位期望。

这个情形下，是否鸣牌的判断也使用一样的蒙特卡洛模拟，分别计算鸣牌和不鸣牌的平均顺位期望。