

# 多智能体系统作业： 传统棋牌游戏的决策方案设计（麻将）

助教：张钊为

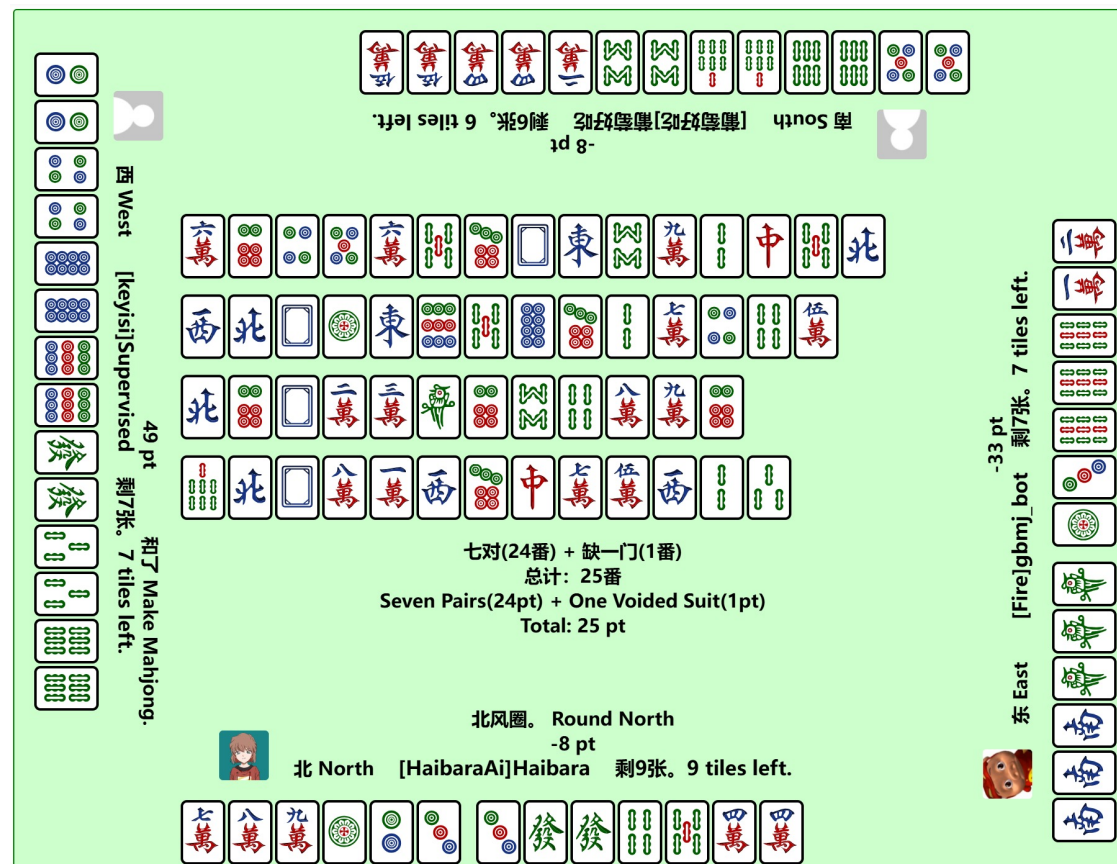
北京大学人工智能研究院

邮箱： [zwzhang@stu.pku.edu.cn](mailto:zwzhang@stu.pku.edu.cn)

微信：tomdy1117

# 实验环境

## Botzone 国标麻将 (Chinese-Standard-Mahjong)



Botzone国标麻将支持人类、AI对战及对局可视化，可以上传模型及算法进行在线对战

在线平台: <https://botzone.org.cn>

游戏信息: <https://wiki.botzone.org.cn/index.php?title=Chinese-Standard-Mahjong>

Botzone交互方法: <https://wiki.botzone.org.cn/index.php?title=Bot>

课程团队同时提供本地强化学习环境供采样训练使用

# 实验环境

## Botzone 国标麻将 (Chinese-Standard-Mahjong)

Botzone国标麻将是基于国标麻将游戏开发的在线游戏平台，国标麻将游戏规则参见

<https://wiki.botzone.org.cn/index.php?title=Chinese-Standard-Mahjong>

## 国标麻将强化学习环境

国标麻将强化学习环境包含一个国标麻将本地交互环境和配套的学习框架。环境内实现了基本的PPO算法，适用于国标麻将Agent的并行采样和训练。我们同时提供一个[国标麻将算番库](#)。

国标麻将强化学习环境的其他细节详见环境文档。

## IJCAI 2025 麻将人工智能比赛

我们同时在本学期举办IJCAI 2025 麻将人工智能比赛，欢迎选做国标麻将作业的各位同学积极报名参加。

比赛设置丰厚的奖项及奖金，优胜队伍有机会前往IJCAI 2025会议现场作线下报告和研讨。

比赛报名截止时间：2025年6月9日

比赛网页：[https://botzone.org.cn/static/gamecontest2025a\\_cn.html](https://botzone.org.cn/static/gamecontest2025a_cn.html)

比赛通知：[https://mp.weixin.qq.com/s/ZjmlIix\\_2xnZUfT9woON3w](https://mp.weixin.qq.com/s/ZjmlIix_2xnZUfT9woON3w)

同学们也可以在比赛网页获得国标麻将强AI的对局数据供辅助训练和学习。

# ■ 具体内容

- 目标：熟悉多智能体强化学习的算法框架和训练流程，基于传统棋牌游戏规则自主设计解决方案，综合运用强化学习和启发式搜索、特征工程等多种方法提升智能体的决策能力。
- 作业任务与要求：
  - 一、（必做）智能体开发，开发基于强化学习的适用于Botzone国标麻将游戏的智能体。设计算法的过程中可以考虑的角度包括但不限于：
    - ◆ **特征处理**。本地交互环境提供了基本的游戏信息和特征表示，开发过程中，也可以基于人类经验自行探索和设计一些特征和特征结构。良好的特征提取和表示有助于帮助模型理解局面状态和决策。
    - ◆ **模型调优**。尝试不同的模型结构并比较不同模型的表现，选择最优的模型进行决策或在决策时参考多个模型输出的行动。由于Botzone存在存储空间和交互时间的限制，请注意控制模型的参数量。
    - ◆ **算法创新**。在强化学习的基础上改进算法，可以融合其他的基于规则、搜索或学习的方法来提升智能体的表现，也可以使用其他方法（见参考文献[1][2]）。
  - 二、（选做）开放式探究，基于Botzone国标麻将平台、本地环境和智能体，对感兴趣的多智能体博弈问题做探究，也可以比较分析多种学习方法的性能和表现。

# 提交说明

- 组队形式：3人一组（需说明每人具体贡献）
- 最终提交内容需将以下几部分压缩打包命名为“学号+姓名”，提交截至时间为第16周周一晚上12点
  - 1. **Botzone智能体**，在Botzone提交国标麻将bot并参加对战评测。
  - 2. **研究性报告**，pdf形式，中英文皆可，页数建议不少于3页，内容包括但不限于算法细节、优化和创新点、开放式问题研究方案、有趣的现象和启发等，必要时可配图说明，同时注意参考文献的引用。
    - 1) Latex版采用ICML2022模板  
(<https://media.icml.cc/Conferences/ICML2022/Styles/icml2022.zip>)
    - 2) Word版无给定模板，但总体要求采用双栏，单倍行距。  
  
字体：英文为Time New Roman，中文为宋体。  
  
字体大小：总标题为三号，章节标题为四号，其他内容为五号。
  - 3. **完整代码和模型**，基于Python3.8+、Pytorch1.4+编写，说明具体安装流程和脚本执行方法，并保证可在ubuntu18+上运行成功，训练好的策略网络模型需保存为.pt文件一并提交。
- 评分角度：综合考虑智能体决策水平、决策算法的创新点、代码性能、和写作展示。

# ■ 评分说明

---

一、只要各组提交的智能体能够正常运行且评测得分超过基线 Bot，代码完整提交、过程符合学术诚信标准就可以获得大量基础分。

二、鼓励锻炼学术和创新性探索，不注重模型绝对性能。

# 参考文献

---

[1] Ge Zhenxing, Xiang Shuai, Tian Pinzhuo, Gao Yang. Solving GuanDan Poker Games with Deep Reinforcement Learning[J]. Journal of Computer Research and Development, 2024, 61(1): 145-155. DOI: [10.7544/issn1000-1239.202220697](https://doi.org/10.7544/issn1000-1239.202220697)

[2] Guan Yang, Minghuan Liu, Weijun Hong, Weinan Zhang, Fei Fang, Guangjun Zeng, Yue Lin. PerfectDou: Dominating DouDizhu with Perfect Information Distillation. NeurIPS 2022. DOI: [10.48550/arXiv.2203.16406](https://doi.org/10.48550/arXiv.2203.16406)