



# Snakes -3v3比赛经验介绍

---

USTC MCCLAB

赵鉴 薛东昀 李子涵

20210820



# 目录

---

- 团队介绍
- 项目背景
- 解决方案
- 反思与展望

# 团队介绍

- 中科大多媒体计算与通信教育部-微软重点实验室
- 队长
  - 赵鉴：博士研究生
  - 薛东昀、李子涵：大二新生
- 指导老师
  - 周文罡、李厚强老师
- 欢迎大家报考研究生
  - zhwg@ustc.edu.cn
  - lihq@ustc.edu.cn



# 游戏+AI开启人工智能的热潮

## ● 业界对游戏AI的探索



游戏模拟环境作为人工智能的研究验证环境，已形成广泛共识

## ● MCC Lab对游戏AI的探索



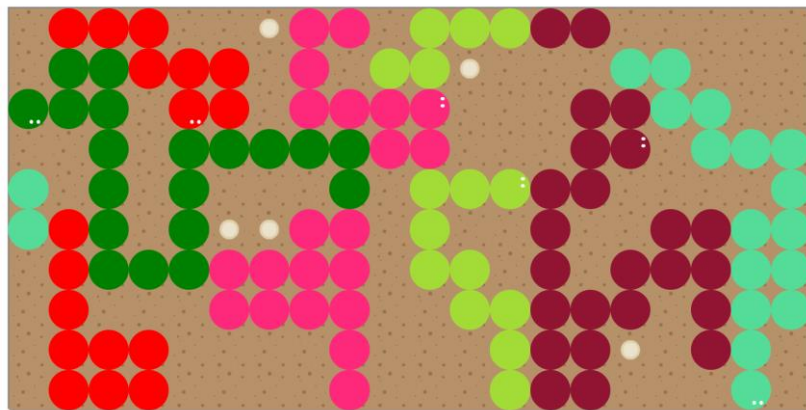
# 比赛背景

## □ 游戏简介

- 贪吃蛇通过控制蛇头方向吃豆子，从而使得蛇变得越来越长。
- 每个玩家控制的蛇的数量为3，每回合玩家同时做出决策控制自己的蛇。

## □ 比赛难点：

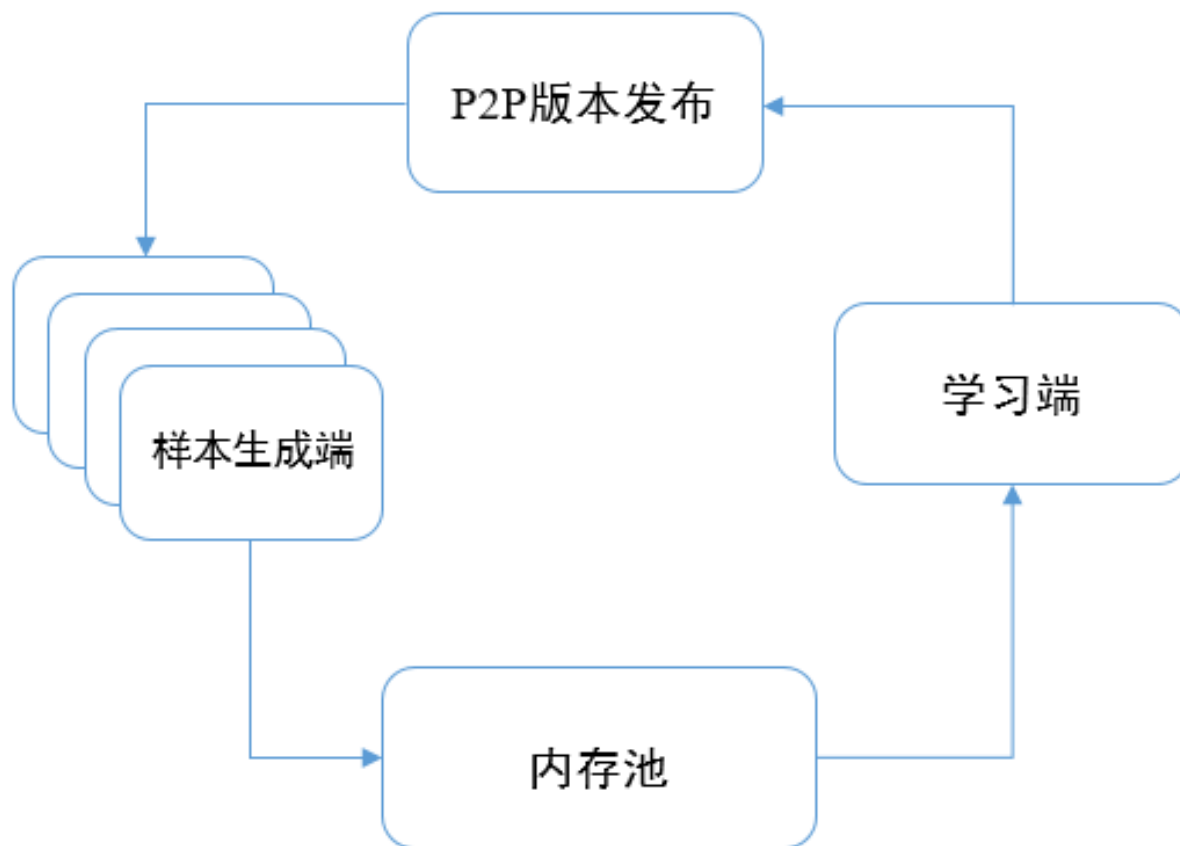
- 获胜条件为蛇身长度总和更长，需要考虑己方三条蛇的合作。
- 蛇头超过边界时可以**穿越**到对称位置，需要对地图特征进行特定处理。
- 当对手水平提高时，地图中可走的位置会越来越少，因此需要产生更具有**进攻性**的AI才能保证获胜。





# 解决方案

- 采用自博弈+强化学习设计框架





# 解决方案

## □ 特征设计

- 采用地图 $10 \times 20 \times 11$  one-hot 矩阵式特征，以捕捉位置关系。

1. 所有蛇的全身位置
2. 此时控制的蛇的头的位罝
3. 我方所有蛇头的位置
4. 敌方所有蛇头的位置
5. 所有豆的位置
6. 此时控制蛇的全身位置
- 7-8. 我方另外两条蛇全身位置
- 9-11. 敌方三条蛇全身位置

## □ 网络设计

- 采用残差网络Resnet提取特征

- circular padding

- ✓ 贪吃蛇越过边界后会从对侧出来，所以可认为整个地图是循环的
- ✓ 对所有的卷积与池化层使用circular padding, 即用对侧的值做padding。



# 解决方案

## □ 零和奖励设计

- 设计单条蛇的零和奖励：

$$\hat{r}_i = r_i - \bar{r}'$$

- ✓ 上式中 $r_i$ 为第 $i$ 条蛇与上一帧的长度差， $\bar{r}'$ 为敌方3条蛇长度差的平均。

- 为鼓励团队配合，奖励设计为：

$$r_i^* = \hat{r}_i \times (1 - \alpha) + \bar{\hat{r}} \times \alpha$$

- ✓ 上式中 $\alpha$ 为超参团队奖励系数， $\bar{\hat{r}}$ 为团队零和奖励的平均。

- 加入赢局奖励 `final_reward=10`

- ✓ 赢得比赛获得奖励，输则给予负奖励

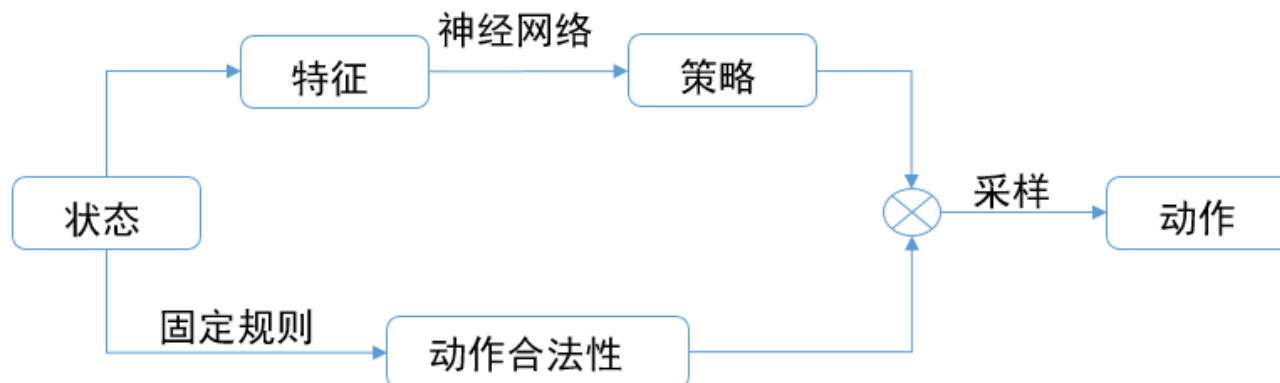




# 解决方案

## □ 提高训练效率

- 各智能体共享网络参数
- 通过固定规则减小动作空间:
  - ✓ 根据当前状态，禁用毫无意义或会造成不良影响的动作
  - ✓ 不允许蛇前往目前蛇头周围4格中有蛇身体的格；发现若此格是某蛇的尾则可以前往。





# 反思与展望

## □ 特征设计

- 以向量形式加入各蛇长度信息
  - ✓ 希望训练出“同归于尽”的策略
  - ✓ 帮助模型了解各条蛇重要程度
- 加入合法动作信息

## □ 网络设计

- 使用中心化训练，促进各蛇之间的配合。
- 使用model-based方法

THANKS!

周文罡老师邮箱: [zhwg@ustc.edu.cn](mailto:zhwg@ustc.edu.cn)

李厚强老师邮箱: [lihq@ustc.edu.cn](mailto:lihq@ustc.edu.cn)