

# 马琰

(+86) 152-6982-9096 • 20210860024@fudan.edu.cn • GitHub • Zhihu

## 教育

复旦大学, 上海, 中国

09/2020 – 06/2023 (预期)

计算机应用技术 硕士, GPA:3.62/4

大连理工大学, 大连, 中国

09/2016 – 06/2020

计算机科学与技术 学士, GPA: 4.076/5, 排名: 12/123

## 项目

**Soccer AI Imitation:** 特定进球风格的足球 AI 模仿学习 (网易互娱 AI Lab 实习) 07/2022 – 09/2022

- 对 COG 2022 5v5 足球比赛中具有特定进球风格的策略进行模仿, 提升己方面对特定对手时的对抗能力
- 针对少量的比赛数据 (30+ 场), 采用生成对抗模仿学习 (GAIL), GAIL 策略依次控制球场上的每个球员
- 设计表征特定进球风格的状态特征, 实现了对“传球式进球”风格的模仿

**Diverse Open Loop Motion Control:** 基于动作序列隐空间的多样开环运动控制 06/2022 – 07/2022

- 从生成任务的角度审视基于多样性的运动控制问题, 尝试学习控制器的隐空间表示
- 通过轨迹生成器产生动作序列控制器, 以此为基础构建先验分布, 从中抽样动作序列以供隐空间的学习
- 通过变分自编码器捕获动作序列的隐空间分布, 隐空间编码了控制器核心特征 (e.g. 方向和速度), 从隐空间采样并通过解码器生成控制器, 控制智能体展现多样的运动模式

**reRLs:** 主流强化学习算法的复现

01/2022 – 03/2022

- 强化学习算法复现。目前已经实现的算法: VPG, A2C, NPG, TRPO, PPO
- 支持基于 Ray 的并行采样和基于 EnvPool 的向量环境, 以及基于Aim的实验记录

## 论文

**Yan Ma, Tianxing Liu, Bingsheng Wei, Yi Liu, Kang, Xu and Wei Li, “Evolutionary Action Selection for Gradient-based Policy Learning”, *International Conference on Neural Information Processing (ICONIP)*’2022, [PDF]**

- 演化强化学习 (Evolutionary Reinforcement Learning) 的方法中普遍存在的一个问题, 即利用进化算法去直接在策略网络的参数空间进行寻优。这会导致混合方法中进化部分难以搜索到良好的策略网络参数
- 本文提出了 Evolutinoary Action Selection (EAS), EAS 的核心思想是将混合方法中进化部分的优化目标从高维难优化的参数空间转移到低维易优化的动作空间
- 将 EAS 与 RL 算法结合为 EAS-RL。RL 策略根据当前状态输出动作, EAS 将动作进化以促进 RL 策略学习。EAS-TD3 在密集奖励和稀疏奖励的运动控制任务中均有着良好的性能表现

**Kang Xu, Yan Ma, Wei Li, Open-ended Diverse Solution Discovery with Regulated Behavior Patterns for Efficient Cross-Domain Adaptation, *Under Review***

- 多样性强化学习 (Diversity-driven Reinforcement Learning) 的方法在训练多样化的策略时通常没有任何的行为约束, 这会导致策略在跨域迁移 (Cross-Domain Adaptation) 中的低效性
- 本文提出了 Diversity in Regulation (DiR), 旨在训练具有约束行为模式的多样化政策以提升的跨域迁移能力。DiR 利用带有状态掩码的逆动力学模型作为多样性度量, 通过定制状态掩码以控制智能体在特定的行为模式下展现多样性
- DiR 在电机故障、传感器故障、动力学参数偏移等跨域测试环境中均展现出了良好的迁移效果

## 获奖

- 复旦大学硕士生优秀学业奖学金 2021
- 大连理工大学优秀毕业生 2020
- 大连理工大学学习一等奖学金 2019

## 技能

编程语言和工具: Python, C/C++, Bash, Numpy, PyTorch, NeoVim, Tmux, Ray, Git, L<sup>A</sup>T<sub>E</sub>X