# 多智能体系统协作定位与目标跟踪

# 基本信息

## 1.背景意义

自然界有很多这样由上百个体构成的大型动物群体，例如鸟群、鱼群，而人类的社会属性也带来了复杂的社交网路关系，集群现象和问题是研究的一个趋势，随着单智能体的强化，研究热点逐步发展到多智能体上，根据具体应用中智能体之间的关系，将多智能体问题分为完全合作式、完全竞争式、混合关系式三种类型，未来会产生更加成熟集群系统，如表演无人机编队，无人货/汽车网络，微型机器人集群，以及游戏虚拟现实以及自动驾驶等。

## 2.简介

多智能体的是复杂大规模多智能体系统（Large-scale Multi-Agent Systems，LMAS），强化学习则一般不需要对系统建立模型，是一种“试错”的控制方法。多智能体系统传统的方法是基于对多智能体系统进行微分方程建立控制模型，进而针对模型进行控制；而近年来强化学习在微型和小规模的多智能体问题中得到了广泛应用，多体强化学习也称之为MASL，基于单智能体强化学习的发展，随着对多智能体问题的深入，多智能体问题形式丰更加富。

## 3国内外现状

中国科学技术大学计算机科学与技术学院的副教授吴锋，主要研究方向为多智能体系统、自动规划理论和智能机器人决策；腾讯AI Lab最近开发的王者荣耀AI——“觉悟”使用了MARL（Multi-Agent Reinforcement Learning）技术，打败了职业选手。DeepMind也使用此技术开发了一个对战小游戏的AI，效果不错。UCL[汪军](https://www.zhihu.com/search?q=%E6%B1%AA%E5%86%9B&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra={"sourceType":"answer","sourceId":794945188}" \t "_blank)教授团队有些很好的成果，也有相关应用。

Neural-Fly 是与加州理工学院计算与数学科学教授 Yisong 开发的，它基于早期的[神经群](https://www.caltech.edu/about/news/machine-learning-helps-robot-swarms-coordinate)系统。Neural-Swarm 教无人机在彼此靠近的情况下自主飞行。使用 GLAS（[Global-to-Local Safe Autonomy Synthesis](https://resolver.caltech.edu/CaltechAUTHORS:20200514-141356088) ） 和 Neural-Swarm 时，机器人不需要对其正在通过的环境或其他机器人打算采取的路径进行完整而全面的了解。机器人学习如何在飞行中导航空间与其局部环境的信息。由进行分散计算，在进入“学习模型”时将新信息整合到运动中

## 4.难点

多智能体在应用的过程中，有很多困难，有以下几点：

一致收敛性与好奇心机制 （稳定性分析和有向无向探索）；

状态空间计算性能（特征工程简化） ；

影响因素与奖励因子的策略设计（网络设计）

网络用基于若干动机指标（motivational indices）的评分函数对观察到的非己环境实体给出评分，并由此排序、裁剪、聚合以提取特征用于决策。第二，网络有一个独特的驱动子网络来引导和训练评分函数的评价倾向和网络参数，确保评分函数能深度绑定所给动机指标。在网络的基础上，设计策略梯度模型，有些问题的核心在于值函数分解（credit assignment），如SMAC。另外一些问题侧重于较为困难合作形式的形成（emergence）。

# 案例和演示

## 2.1关于多智能体的一致性仿真**https://github.com/Bluefeicai/MultiAgentSystem**

## 2.2一阶二阶多智能体一致性控制的Matlab程序

<https://github.com/Bluefeicai/MultiAgentSystem>

# 研究内容

## 3.1博弈与优化的多智能体

多智能体强化学习算法的理解分析。包括线性规划求解双智能体零和博弈，矩阵博弈（每个智能体的策略是一个关于其动作空间的概率分布，每个智能体的目标是最大化其获得的奖励值），两个智能体的矩阵博弈中的纳什均衡（一个所有智能体的联结策略），随机博弈（是马尔可夫决策过程与矩阵博弈的结合）。

**MARL与博弈论的参考介绍** <https://zhuanlan.zhihu.com/p/383112880>

## 3.2图网络强化学习与多智能体路径规划（GNN & Multi-Robot Path Planning）

没有智能体之间的通信信息,难以实现智能体之间的合作和全局最优的，每一个agent可当作图中的一个节点,两个节点之间有没有联系可取决于他们之间的距离等, 最后由建立GNN图神经网络汇聚信息。

3.3算法综述：一些论文研究的算法有（摘取网络）  
1、ROMA: Multi-Agent Reinforcement Learning with Emergent Roles. Tonghan Wang, Heng Dong, Victor Lesser, Chongjie Zhang.  
2、DOP: Off-Policy Multi-Agent Decomposed Policy Gradients. Yihan Wang, Beining Han, Tonghan Wang, Heng Dong, Chongjie Zhang  
3、QPLEX: Duplex Dueling Multi-Agent Q-Learning. Jianhao Wang, Zhizhou Ren, Terry Liu, Yang Yu, Chongjie Zhang  
4、LICA: Learning Implicit Credit Assignment for Cooperative Multi-Agent Reinforcement Learning. Meng Zhou, Ziyu Liu, Pengwei Sui, Yixuan Li, Yuk Ying Chung  
5、Value-Decomposition Multi-Agent Actor-Critics. Jianyu Su, Stephen Adams, Peter A. Beling.  
6、Tesseract: Tensorised Actors for Multi-Agent Reinforcement Learning. Anuj Mahajan, Mikayel Samvelyan, Lei Mao, Viktor Makoviychuk, Animesh Garg, Jean Kossaifi, Shimon Whites**·**n, Yuke Zhu, Animashree Anandkumar.

1. SMIX(λ): Enhancing Centralized Value Functions for Cooperative Multi-Agent Reinforcement Learning. Xinghu Yao, Chao Wen, Yuhui Wang, Xiaoyang Tan  
   网友报告的  
   8、NDQ: Learning Nearly Decomposable Value Functions with Communication Minimization. Tonghan Wang\*, Jianhao Wang\*, Chongyi Zheng, Chongjie Zhang  
   9、Graph convolutional reinforcement learning. Jiechuan Jiang, Chen Dun, Tiejun Huang, Zongqing Lu  
   10、ATOC: Learning Attentional Communication for Multi-Agent Cooperation. Jiechuan Jiang Peking University Zongqing Lu  
   11、[UPDeT: Universal Multi-agent RL via Policy Decoupling with Transformers](https://zhuanlan.zhihu.com/p/408515796/edit)[Siyi Hu](https://link.zhihu.com/?target=https://arxiv.org/search/cs?searchtype=author&query=Hu%2C+S" \t "https://zhuanlan.zhihu.com/p/_blank),[Fengda Zhu](https://link.zhihu.com/?target=https://arxiv.org/search/cs?searchtype=author&query=Zhu%2C+F" \t "https://zhuanlan.zhihu.com/p/_blank),[Xiaojun Chang](https://link.zhihu.com/?target=https://arxiv.org/search/cs?searchtype=author&query=Chang%2C+X" \t "https://zhuanlan.zhihu.com/p/_blank),[Xiaodan Liang](https://link.zhihu.com/?target=https://arxiv.org/search/cs?searchtype=author&query=Liang%2C+X" \t "https://zhuanlan.zhihu.com/p/_blank)  
   12、FOP: Factorizing Optimal Joint Policy of Maximum-Entropy Multi-Agent Reinforcement Learning

13、[FACMAC: Factored Multi-Agent Centralised Policy Gradients](https://link.zhihu.com/?target=https://arxiv.org/abs/2003.06709" \t "https://zhuanlan.zhihu.com/p/_blank)

14、Learning Individual Intrinsic Reward in MARL

# 四．涉及基础

## 0.入门学习

**算法学习**

**教程1：**<https://zhuanlan.zhihu.com/p/53811876>

**（博弈，[基础算法（MiniMax-Q，NashQ，FFQ，WoLF-PHC）](https://zhuanlan.zhihu.com/p/53563792" \t "https://www.zhihu.com/people/kai-shi-xiong/_blank)[分布式学习算法](https://zhuanlan.zhihu.com/p/53622102" \t "https://www.zhihu.com/people/kai-shi-xiong/_blank)，[MADDPG算法](https://zhuanlan.zhihu.com/p/53811876" \t "https://www.zhihu.com/people/kai-shi-xiong/_blank)，[DRQN](https://zhuanlan.zhihu.com/p/54898904" \t "https://www.zhihu.com/people/kai-shi-xiong/_blank)，[QMIX，](https://zhuanlan.zhihu.com/p/55003734" \t "https://www.zhihu.com/people/kai-shi-xiong/_blank)[AC for CDec-POMDP，](https://zhuanlan.zhihu.com/p/66571753" \t "https://www.zhihu.com/people/kai-shi-xiong/_blank)[图卷积强化学习DGN](https://zhuanlan.zhihu.com/p/143816617" \t "https://www.zhihu.com/people/kai-shi-xiong/_blank)，[注意力图网络的多智能体博弈](https://zhuanlan.zhihu.com/p/146460071" \t "https://www.zhihu.com/people/kai-shi-xiong/_blank) ）**

**参考教程2：<https://www.zhihu.com/people/zhiqianghe/posts> 东北大学学生知乎号**

**强化学习入门知识（基础篇）**

<https://zhuanlan.zhihu.com/p/521531331>

<https://www.zhihu.com/people/kai-shi-xiong/posts>

**路径规划例程上手**

**SARSA** <https://blog.csdn.net/SayHello2You/article/details/122209228?spm=1001.2014.3001.5502>

**A\*+RL** <https://zhuanlan.zhihu.com/p/337729123>

## 1.数学基础

[线性已上过需要继续课程：代数](https://www.zhihu.com/search?q=%E7%BA%BF%E6%80%A7%E4%BB%A3%E6%95%B0&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra={"sourceType":"answer","sourceId":"924692930"}" \t "https://www.zhihu.com/_blank)、概率，数值分析，运筹优化；

需要了解学习：微分方程（matlab差分方程），几何，多元统计，、图论等

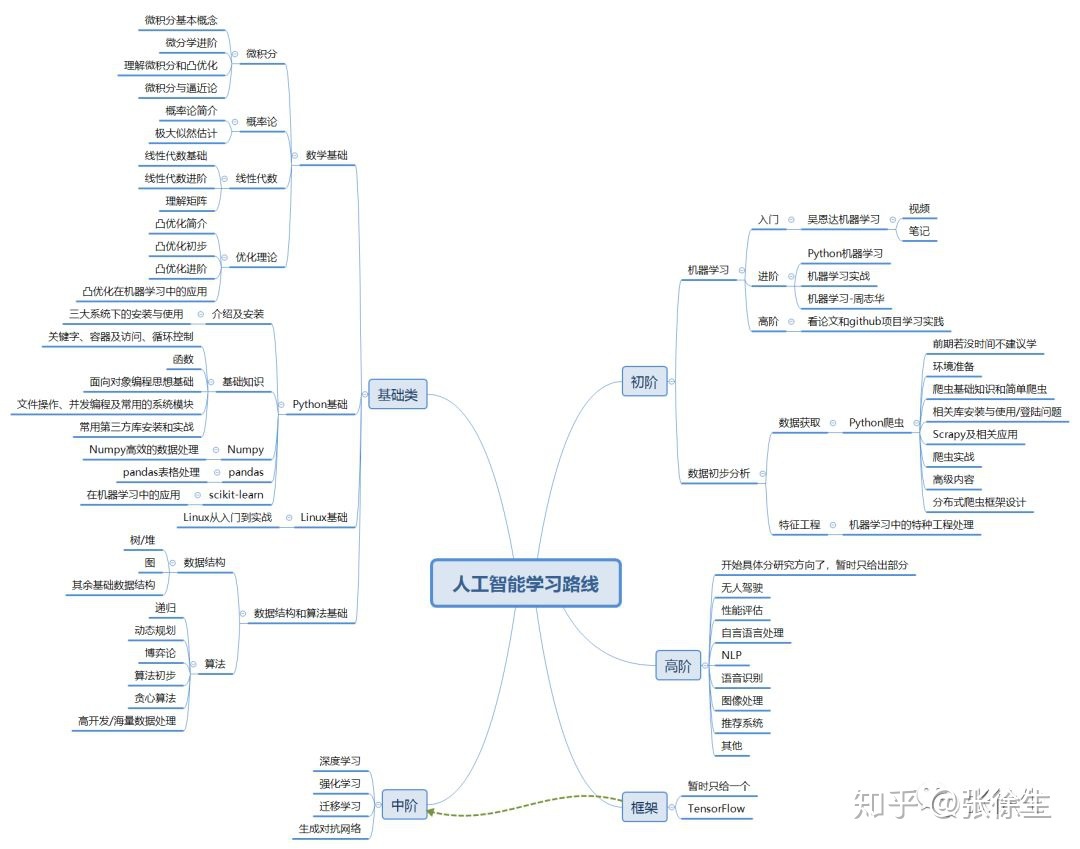
## 2.计算机基础

编程：C（matlab），python（AI）

通信：计算机网络，通信协议，信息论。

控制 ：计算机控制理论，现代控制理论，智能控制基础，模式识别

AI：机器学习，深度学习，强化学习；



## 3.Masl基础

## 

相关书籍： https://github.com/chengji253/Formation\_control\_books

入门图文教程 <https://www.cnblogs.com/ting1/p/16833963.html> 定位相关 高翔的《[视觉SLAM14讲](https://www.zhihu.com/search?q=%E8%A7%86%E8%A7%89SLAM14%E8%AE%B2&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra={"sourceType":"article","sourceId":"375078582"}" \t "https://www.zhihu.com/_blank)》

[编队协同方向](https://www.zhihu.com/search?q=%E7%BC%96%E9%98%9F%E5%8D%8F%E5%90%8C%E6%96%B9%E5%90%91&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra={"sourceType":"article","sourceId":"375078582"}" \t "https://www.zhihu.com/_blank)： 《**Formation control**》 Hyo-Sung Ahn 韩国 大佬有Matlab示例代码

# 排期计划

“预报毕业的同学请注意，23届毕设相关时间安排如下： 学生选题时间：10月31日 - 11月13日 任务书下达时间：11月14日 - 11月26日（10-12周） 开题报告提交截止时间：12月30日（因部分同学考研，截止时间推迟） 中期答辩：下学期开学初第2-3周 任务书是导师下达，开题报告是学生提交后导师审核。”

2022年10月31-11月13日

## 环境配置（）

例如：游戏环境库pip

https://blog.csdn.net/Bit\_Coders/article/details/125544958?spm=1001.2014.3001.5506

## 例程运行了解（）

11月13日-11月27日

## 基础学习（要点）

**11月27日 -12月18日 （中期答辩）**

## 论文学习（重点）

## 总结分析

2023年1月2日-1月21日

## 复现与掌握

## 仿真

1月29日 -1月22日

## 运用与创新

## 调试。

# 六．附录