

吴文浩

电话: 18005527326 邮箱: wenhaowu@smail.nju.edu.cn

籍贯: 安徽蚌埠市 政治面貌: 共青团员



教育背景

- 南京大学 工程管理学院 控制工程 2024.09 至今
- 南京大学 工程管理学院 自动化 2020.09~2024.06
- 获奖情况: 人民奖学金 (2 次)、硕士一等学业奖学金 (2 次)、高额奖学金 (1 次)。
- 主要课程: 深度强化学习、人工智能、机器学习、群智能体的协作理论与方法。

研究经历

- From Conflict to Consensus: Boosting Medical Reasoning via Multi-Round Agentic RAG** [ICML 2026](#), 一作
研究动机: 传统的 single-round RAG 无法动态检索模型推理过程中的知识缺口, 而基于 token-level uncertainty 的 adaptive RAG 方法存在检索信号噪音、模型过度自信导致信号不准确等问题, 限制了检索效率和在医疗场景的应用。
研究方法: 基于多轮迭代改进框架提出 **conflict-guided** 检索方法, 利用模型生成中采样的多条 responses 之间的冲突作为检索信号, 自适应生成精准查询语句; 在融入历史回答后的上下文优化方面, 从 In-Context RL 角度出发, 利用 sequence-level entropy 或轻量化 BERT 验证器对历史轨迹**打分重排序**, 以引导模型生成更高质量的回答。
- Mixture-of-Experts Meets In-Context Reinforcement Learning** [NeurIPS 2025](#), 一作
研究动机: In-context RL 方法输入的轨迹包含三个模态的信息, 即状态、动作、奖励, 以及多样化的任务轨迹。同一模块难以有效提取多模态和多任务信息, 并且存在梯度冲突问题, 限制了多任务学习效率。
研究方法: 提出包含两个并行 MoE 结构的框架。针对输入中的三个模态信息, 提出 **Token-wise MoE**, 附加 **balance loss** 进行优化, 将不同模态的信息分配给不同的专家处理; 针对输入中包含的多任务信息, 提出 **Task-wise MoE**, 利用 **contrastive learning** 更准确提取任务信息, 并将不同任务的轨迹分配给不同专家处理。
- Text-to-Decision Agent: Offline Meta-Reinforcement Learning from Natural Language Supervision** [NeurIPS 2025](#), 共一
研究动机: 现有的智能体决策系统难以实现基于语言指令的泛化决策, 限制了强化学习决策的实际应用。同时, 强化学习的轨迹数据与自然语言之间存在语义鸿沟, 限制了智能体对语言指令的理解和执行能力。
研究方法: 首先通过 **world model encoder-decoder** 的方式训练轨迹编码器, 将轨迹编码为包含环境动态信息的决策嵌入; 然后利用 **contrastive learning** 对预训练语言模型进行 LoRA 微调, 实现决策嵌入和文本嵌入的语义对齐; 最后基于对齐的文本嵌入构建条件化的**可拓展策略框架**, 并在 Decision Diffuser 和 Decision Transformer 上进行验证。
- Meta-DT: Offline Meta-RL as Conditional Sequence Modeling with World Model Disentanglement** [NeurIPS 2024](#), 学生二作
研究动机: 在语言和视觉领域, 基于 Transformer 模型并通过不断扩大数据集规模, 展现出显著的通用智能效果, 即能够从历史经验中学习并且泛化到新任务, 然而强化学习智能体在相同范式下表现出较差的泛化性能。
研究方法: 预训练一个具有上下文感知能力的 **world model**, 学习紧凑的 **task representation**, 作为 Causal Transformer 的提示输入, 指导面向任务的序列生成; 此外, 将策略生成的历史轨迹作为 **self-guided prompt**, 并且利用预训练的世界模型选择产生最大误差的轨迹片段, 作为世界模型信息的 **complementary prompt**。

实习经历

- 华为诺亚方舟实验室** 2025.07~2026.01
实习内容: 将检索增强生成 (RAG) 技术与医疗问答场景深度融合, 缓解大语言模型的“幻觉”问题。具体包括: 搭建医学知识检索服务端; 设计自动检索算法框架, 基于冲突生成精准查询语句; 微调 BERT 模型以适应医疗问答场景, 对模型回答进行打分排序; 结合反思改善机制, 尝试基于支持向量机的自适应终止方法; 探索 Agentic RL 在医疗推理中的应用, 利用强化学习训练模型推理思路和方向选择。
实习成果: 探索 Adaptive RAG 与 Agentic RL 领域前沿课题; 以**第一作者**身份在顶级国际会议 **ICML 2026** 上发表学术论文。