韩槿一

15224975562(微信) jinyihan099@gmail.com https://jinyihan99.github.io

研究方向

- **Language Agents:** 聚焦构建具备自主决策、持续学习与长周期任务执行能力的agents，包括**自进化方法**、**技能沉淀与复用**、agentic RL及 Agent 评测等核心技术
- **Reasoning Models:** 增强大模型的复杂推理能力以及**工具使用能力**，以及推理效率优化
- **Post-Training:** 探索大模型**持续自进化**、课程学习策略及强化学习优化方法，以提升模型性能

教育经历

华东师范大学 (复旦大学联合培养) 学术博士 计算机科学与技术学院 2023.09 – 2027.06
东华大学 学术硕士 计算机科学与技术学院 2020.09 – 2023.03
河南财经政法大学 本科 计算机科学与技术学院 2016.09 – 2020.06

实习经历

腾讯 TEG 混元大模型团队 **青云计划** 2026.05 – 至 今

LLM Agent 技能遵循能力评测基准

- **目标:** 首次提出面向 Agent 场景的**过程合规性**评测范式，评估模型是否按领域技能规范（Skill）的步骤顺序、工具选择、行为边界等过程约束执行，而非仅关注任务完成度。
- **内容:** 设计七阶段自动化合成流水线，构建覆盖 9 大领域的 **90 个 Skill**、**200+ Task**、**400+ 评分规则**，对 7 个主流模型系统评测。揭示当前最强模型过程遵循率仅 ~64%，开源模型普遍存在"任务能做对但不按规矩做"的遵循鸿沟，为 Agent 可控性优化提供量化基线与诊断依据。

字节跳动 中国交易与广告 **筋斗云人才计划**（多模态大模型算法工程师） 2025.07 – 2026.04

01 - Reasoning-based SafetyGPT-GRM 奖励模型训练

- **背景:** 面向电商治理与安全场景，研发具备**推理能力**的通用奖励模型（GRM），用于对前置模型 CoT 推理进行**反思校验**、**质量过滤与可解释分析**。
- **内容:** (1) 提出推理判定范式 $PRA \to A'R'A''$，通过**首 token** A' 实现高置信度快速筛选，推理部分 R' 提供可解释反思；(2) 设计 **SFT + RL Merge Loss** 联合训练，首 token 用 SFT Loss 保证判定稳定性，推理部分用 **GRPO Loss** 优化推理质量。
- **效果:** 反思校验相比直接生成准确率提升 **+6.3~6.75 PP**；首 token 高精度筛选 Recall 达 **26.86%**（vs A'' 提升 +26.77 PP）；90% Precision 下 Merge Loss 相比纯 GRPO Recall 提升 **+2.9 PP**。

02 - 仲裁智能化算法（客服专项）

- **背景:** 面向电商仲裁场景，在申诉率不提升、满意度不下降前提下替代部分人工客服。
- **内容:** (1) 参与设计并落地 **SOP 规则驱动的模块化多模型仲裁系统**，构建**标签识别** → **SOP 检索** → **时效判断** → **举证判定** → **售后方案生成**全自动化流程；(2) 训练并上线 policy 检索、时效判断、判证、方案生成等多个专项模型。
- **效果:** 核心模型已**上线应用于仲裁系统**，领域准确率 85%+，相较豆包等模型提升近 **10 PP**。

复旦大学 知识工场实验室 **Research Intern** 2023.09 – 2025.06

- 主导多个 LLM Reasoning & Agent 方向研究项目，以一作/二作在**ICLR**、**ACL**、**ICML**、**AAAI**发表10+篇论文，并指导近**20名**硕士及本科生完成科研合作与论文产出。

项目经历

01 - GenericAgent GitHub (13K+ Stars) Technical Report **贡献人** 2026.01 – 至 今

- 以"上下文信息密度最大化"为核心理念来设计自进化 LLM Agent 系统，已开源并发布技术报告，社区用户达4K+。
- 设计并参与三阶段自进化流水线（自然语言执行 → SOP 蒸馏 → 代码化执行），将成功 trajectory 自动沉淀为可复用 SOP 与可执行代码，全流程无需人工介入，使用沉淀的skill，跨8项任务token平均节省 79.3%

02 - 基于深度思考的大模型自主工具学习方法 项目介绍 **主要贡献人** 2025.04 – 2025.09

- 提出基于 **深度思考**的主动工具使用框架，设计 **Think-Act** 与 **Planner-Executor** 两种工具调用机制，并通过强化学习训练模型动态选择工具辅助复杂任务。
- 在 **9 位数复杂数学任务**上，相比 **Qwen2.5-7B-Instruct** 准确率提升 **+16.7 PP**，并在未见任务上展现良好泛化能力。

03 - R1-Zero 深度思考能力高效复现（GRPO） 项目介绍 **主要贡献人** 1.6K Stars 2025.03 – 2025.04

- 开源社区 **最早复现 R1-Zero 自发反思能力**的项目之一，基于 **GRPO** 强化学习训练深度思考能力，核心实现仅约 **200 行代码**。
- 显著降低复现成本：仅 **5 步 RL 训练**即可使 **Qwen2.5-3B** 准确率稳定在 **60%+**（峰值约 **70%**），并在 **200 步内出现"Aha Moment"**。

论文列表

[*: 导师第一作者,学生为第二作者; †: 通讯作者；横线加粗：本人; 曾用名：Haixia Han]

01 - Language Agents

- **GenericAgent: A Token-Efficient Self-Evolving LLM Agent via Contextual Information Density Maximization**
  Jiaqing Liang, **Jinyi Han**, Weijia Li, Xinyi Wang, Zhoujia Zhang, Zishang Jiang, et.al
  2026, Preprint.
- **From Instruction Following to Skill Following: A Process-Oriented Benchmark for LLM Agents**
  Jinyi Han, Yuanjian Xu, Ying Liao, Zhichao Hu, Yanghua Xiao.
  2026. (Under Review)
- **From Outcomes to Actions: Leveraging Hindsight for Long-Horizon Language Agent Training**
  Zishang Jiang, Tingyun Li, **Jinyi Han**, Xinyi Wang, Sihang Jiang, Yizhou Ying, Xiaojun Meng, Jiansheng Wei, Jiaqing Liang, Yanghua Xiao
  2026, ICML.
- **SOPFlow: Compiling Standard Operating Procedures into Repairable Executable Workflows**
  Ying Liao, Ying Huang, **Jinyi Han**, Siyuan He, Jiaqing Liang, Yanghua Xiao
  2026, ACL. (Under Review)

02 - Reasoning Models

- **A Stitch in Time Saves Nine: Proactive Self-Refinement for Language Models**
  Jinyi Han, Xinyi Wang, Haiquan Zhao, Tingyun Li, Zishang Jiang, Sihang Jiang, Jiaqing Liang, Xin Lin, Weikang Zhou, Zeye Sun, Fei Yu, Yanghua Xiao
  2026, ICLR.
- **Your Models Have Thought Enough: Training Large Reasoning Models to Stop Overthinking**
  Jinyi Han, Ying Huang, Ying Liao, Haiquan Zhao, Zishang Jiang, Xinyi Wang, Xikun Lu, Guanghao Zhou, Sihang Jiang, Jiaqing Liang, Weikang Zhou, Zeye Sun, Fei Yu, Yanghua Xiao
  2026, ICLR.
- **Structured Reasoning for Large Language Models**
  Jinyi Han, Zixiang Di, Zishang Jiang, Ying Liao, Jiaqing Liang, Jie Wang, Zheming Yang, Zhi Li, Yongqi Wang, Xiaofeng Ji, Yanghua Xiao, Miao Liu
  2026. (Under Review, Work Done in ByteDance Internship)
- **Small Language Model Can Self-correct**
  Haixia Han, Jiaqing Liang, Jie Shi, Qianyu He, Yanghua Xiao
  2024, AAAI.
- **Think Thrice Before You Act: Progressive Thought Refinement in Large Language Models**
  Chengyu Du, **Jinyi Han**, Yizhou Ying, Aili Chen, Qianyu He, Haokun Zhao, Haoran Guo, Sirui Xia, Jiaqing Liang, Zulong Chen, Liangyue Li, Yanghua Xiao
  2025, ICLR.

- **ADaPT: Token-Level Decoupling for Efficient Large Reasoning Models**
  Tingyun Li, Zishang Jiang, **Jinyi Han**, Xinyi Wang, Sihang Jiang, Han Xia, Zhaoqian Dai, Ma Shuguang, Fei Yu, Jiaqing Liang, Yanghua Xiao
  2026, ACL.

03 - Post-Training

- **Selective Expert Guidance for Effective and Diverse Exploration in Reinforcement Learning of LLMs**
  Zishang Jiang, **Jinyi Han**, Tingyun Li, Xinyi Wang, Sihang Jiang, Zhaoqian Dai, Ma Shuguang, Fei Yu, Jiaqing Liang, Yanghua Xiao
  2026, ICLR.
- **CDS: Data Synthesis Method Guided by Cognitive Diagnosis Theory**
  Haokun Zhao, **Jinyi Han**, Jiaqing Liang, Xiaojun Meng, Jiansheng Wei, Yanghua Xiao
  2025, ACL.
- **Difficulty Is Not Enough: Curriculum Learning for LLMs Fine-tuning Must Consider Utility**
  Zishang Jiang, **Jinyi Han**, Tingyun Li, Xinyi Wang, Sihang Jiang, Jiaqing Liang, Xiaojun Meng, Jiansheng Wei, Yanghua Xiao
  2025, AAAI (Oral).
- **Not All Negative Samples Are Equal: LLMs Learn Better from Plausible Reasoning**
  Zixiang Di, **Jinyi Han**, Shuo Zhang, Ying Liao, Zhi Li, Miao Liu, Xiaofeng Ji, Yongqi Wang, Zheming Yang, Ming Gao, Bingdong Li, Jie Wang
  2026 (Under review, **Equal Contribution**, Work Done in ByteDance Internship)
- **Don't Tell the Answer, Truly Guide the Reasoning During RL Rollouts**
  Xinyi Wang, **Jinyi Han**, Zishang Jiang, Tingyun Li, Jiaqing Liang, Sihang Jiang, Zhaoqian Dai, Ma Shuguang, Fei Yu, Yanghua Xiao
  2026, ACL.
- **Knowing How Certain It Is: Confidence Estimation Throughout LLM Generation**
  Jinyi Han, Tingyun Li, Shisong Chen, Xinyi Wang, Jiaqing Liang, Yanghua Xiao
  2026, AI. (Under Review)

获奖情况

华为火花奖	2024.12、2025.12
上海市优秀毕业生 (Top 5%)	2023.03
东华大学优秀硕士毕业论文 (Top 3%)	2023.03
河南省优秀学位论文 (Top 3%)	2021.07
学业奖学金一等奖	2020.10–2022.10

其他经历

- **校园基金:**
  - **面向大语言模型的持续学习方法研究** 优秀博士生学术创新能力提升计划 2024.09 – 至 今
    项目负责人 已结项
  - **基于大模型的智能教材问答关键技术研究** 人工智能赋能心理/教育学科交叉研究项目 2025.05 – 2025.11
    项目负责人 已结项
  - **面向大语言模型的持续学习方法研究** 博士研究生科研创新基金 2023.12 – 2024.12
    项目负责人 已结项
- **学术服务与合作**：Applied Intelligence 期刊审稿人；ICLR、NeurIPS、ACL、AAAI 等国际会议审稿人；与 Stanford University、Harvard University、University of Toronto 等高校研究者开展学术合作与交流