谭泽霖

(+86) 19392822207 · tanzl@mail.ustc.edu.cn · 微信：MikotoUSTC · Google Scholar

教育背景

中国科学技术大学, 人工智能, *博士研究生* 2025.6 – 2028.6

中国科学技术大学, 计算机技术, *硕士* 2022.9 - 2025.6

中国科学技术大学, 计算机科学与技术, *本科* 2018.9 - 2022.6

实习

上海人工智能实验室：智能体与大模型后训练 Mentor：Chen Zhang and Shuyue Hu 2025.7 – 至今

- **基于自进化 (RSI) 的 AI Infra 优化**：搭建并行搜索的自进化框架，通过基于轨迹的元综合和启发式历史信息管理方式，协同演进 AI Infra 与智能体 Harness。
- **基于多智能体的 Skill 任务数据管线**：提出基于多智能体系统的递归数据合成管线 SKT，**用于合成长交互、Skill 依赖的任务及多轮技能使用轨迹**来训练模型的 Skills 使用能力，并内化部分 harness 能力。
- **大模型智能体的推理范式路由**：提出 Select-then-Solve，在解题前通过轻量级路由器选择 ReAct、Plan-Execute 等推理范式，提升解题准确率与推理效率。
- **基于 Rubric 的大模型强化学习后训练算法**：通过解耦优势归一化，缓解传统 RLVR 的奖励信号衰竭及 Rubric RLVR 中的奖励投机问题。
- **大模型强化学习 Scaling Law**：研究不同模型规模、超参数下，大模型强化学习后训练的扩展规律。
- **AgentPanel 多智能体交互论坛**：参与构建面向科学探索的人机协作多智能体论坛。

美团：大模型推理优化 Mentor：Yineng Zhang and Ke Bao 2024.5-2024.7

- **块粒度自动前缀缓存**：基于块级前缀树管理 GPU 显存，实现大模型推理的自动前缀缓存。

腾讯：大模型推理优化 Mentor:Hande Dong 2023.12-2024.5

- **推理调度策略优化**：在基于 vLLM v0.2.7 的内部引擎中实现动态拆分融合版 Chunked Prefill。
- **SM 感知的算子并行框架**：通过 GPU 卡内资源隔离，并行执行预填充与解码阶段的注意力算子。
- **显存感知的动态分块策略**：按显存余量调整分块大小，缓解高负载下抢占引入的重计算开销。

论文

- SKT: Skill-Use Training at Scale via Verified Synthetic Data Generation（**第一作者，AAAI 2027 在审**）
- PAPO: Stabilizing Rubric Integration Training via Decoupled Advantage Normalization（**第一作者，EMNLP 2026 主会**）
- Scaling Behaviors of LLM Reinforcement Learning Post-Training: An Empirical Study in Mathematical Reasoning（**第一作者，ACL 2026 主会**）
- Sage-A1.5: From Long-Horizon Execution to Scientific Exploration with Capability Composition（**技术报告，2026**）
- AgentPanel: Toward a New Paradigm for Human-AI Collaboration in Exploring Scientific Questions（**DAI 2026 Oral**）
- Scaling the Horizon, Not the Parameters: Reaching Trillion-Parameter Performance with a 35B Agent（**技术报告，2026**）
- Parametric Skills
- Select-then-Solve: Paradigm Routing as Inference-Time Optimization for LLM Agents（**第二作者，ICLR 2027 在审**）
- SkillsBench: Benchmarking How Well Agent Skills Work Across Diverse Tasks（**NeurIPS 2026**）
- AgentAsk: Multi-Agent Systems Need to Ask（**ACL 2026 主会**）
- The Landscape of Agentic Reinforcement Learning for LLMs: A Survey（**TMLR 2025**）

开源社区

- **RL Infra ROLL** Contributor
- AgentPanel Contributor
- **SkillsBench** Contributor
- SM-Aware Operator Parallelism Framework Designer