

Yanjun Chen

PhD Student · Department of Computing · The Hong Kong Polytechnic University

yan-jun.chen@connect.polyu.hk · battam1111.github.io · Google Scholar · github.com/Battam1111 · ORCID 0009-0001-9065-9137

RESEARCH INTERESTS

Environment-Centric AI: treating the training environment as a designed object whose pieces (reward, feedback, observation, evaluation) shape what is learned, across large language models, reinforcement learning, and embodied AI. The longer bet is that environments themselves must *scale*, *train*, and *generalize* the way models have.

EDUCATION

The Hong Kong Polytechnic University, Hong Kong Sept 2024 to present

PhD, Department of Computing.

Advisors: Prof. Wenjie Li (Maggie), Prof. Wei Zhang.

Jinan University, Guangzhou, China Sept 2020 to Jun 2024

B.Eng. in Cyberspace Security. GPA 88.6 / 100 · Rank 6 of 62 · Outstanding Graduate.

PUBLICATIONS

Yanjun Chen = author; ☆ = first author; venues in *italic*.

In submission

1. ☆ **Yanjun Chen**, Yirong Sun, Hanlin Wang, Jinghan Wang, Xinming Zhang, Xiaoyu Shen, Wenjie Li, Wei Zhang. **Exact Is Easier: Credit Assignment for Cooperative LLM Agents**. *arXiv:2603.06859*, 2026.

Peer-reviewed

1. ☆ **Yanjun Chen**, Dawei Zhu, Yirong Sun, Xinghao Chen, Wei Zhang, Xiaoyu Shen. **The Accuracy Paradox in RLHF: When Better Reward Models Don't Yield Better Language Models**. *EMNLP 2024*.
2. Xinghao Chen, Zhixin Sun, Wenjin Guo, Miao Zhang, **Yanjun Chen**, Yirong Sun, Hao Su, Yu Pan, et al. **Unveiling the Key Factors for Distilling Chain-of-Thought Reasoning**. *Findings of ACL 2025*.
3. Yirong Sun, Dawei Zhu, **Yanjun Chen**, Eric Xiao, Xinghao Chen, Xiaoyu Shen. **Fine-Grained and Multi-Dimensional Metrics for Document-Level Machine Translation**. *NAACL 2025*.

Preprints (co-author)

1. Xinghao Chen, Anhao Zhao, Heming Xia, Xuan Lu, Hanlin Wang, **Yanjun Chen**, Wei Zhang, Jian Wang, Wenjie Li, et al. **Reasoning Beyond Language: A Comprehensive Survey on Latent Chain-of-Thought Reasoning**. *arXiv:2505.16782*, 2025.

HONORS & AWARDS

May 2026 PolyU Micro Fund 2025/26 Cohort 2, shortlisted (HK\$20,000 cash prize).

May 2026 HKSTP Ideation Programme, conditional offer.

Jun 2024 Outstanding Graduate, School of Information Science & Technology, Jinan University.

2023 First Prize, 5th Information Mining Competition, Jinan University.

2023 Honorable Mention, The Mathematical Contest in Modeling (MCM).

2022 First Prize, National "Strong Nation Cup" Blockchain Technology Skills Competition.

OPEN SOURCE

AccuracyParadox-RLHF, author / maintainer github.com/Battam1111/AccuracyParadox-RLHF

Official implementation of *The Accuracy Paradox in RLHF* (EMNLP 2024). Provides reproducible training pipelines, evaluation harnesses, and reference reward models for the paper's analysis of the reward-model / language-model quality mismatch.

TECHNICAL SKILLS

Programming Python, C/C++, Bash; Git, Linux.

Frameworks PyTorch, Hugging Face Transformers / TRL / Datasets.

Languages Mandarin (native), English (professional), Japanese (elementary).