YANG XIAO

+(86) 180-2100-2516   ·   23037247r@connect.polyu.hk   ·   [Personal Website](#)   ·   [Google Scholar](#)   ·   [GitHub](#)

EDUCATION

The Hong Kong Polytechnic University   2023 – 2027 (Expected)

- Ph.D. Student in Department of Computing. Advisor: Prof. Wenjie Li and Prof. Pengfei Liu.
- Expected Graduation: September 2027
- **Available for full-time internship throughout 2026, flexible on location**

Fudan University   2018 – 2022

- B.E. in Software Engineering. GPA Rank: Top 1%.

RESEARCH INTERESTS

LLM Post-training, Test-time Reasoning & Scaling, Agentic AI Systems.

Focused on efficient post-training methods for LLMs, with expertise in test-time compute scaling and autonomous agent development. Demonstrated ability to achieve state-of-the-art results with minimal training data (Limo/LIMOPro series).

SELECTED PUBLICATIONS & WORKING PAPERS

1. **Limi: Less is More for Agency**

   - **Yang Xiao**, Mohan Jiang, Jie Sun, Keyu Li, Jifan Lin, Yumin Zhuang, Ji Zeng, Shijie Xia, Qishuo Hua, Xuefeng Li, Xiaojie Cai, Tongyu Wang, Yue Zhang, Liming Liu, Xia Wu, Jinlong Hou, Yuan Cheng, Wenjie Li, Xiang Wang, Dequan Wang, Pengfei Liu.
   - *Under Review, 2026* (**First Author**).
   - Achieved 73.5% on AgencyBench using only 78 samples, 53.7% improvement over 10K-sample baselines.
   - [Code]

2. **AgencyBench: Benchmarking the Frontiers of Autonomous Agents in 1M-Token Real-World Contexts**

   - Keyu Li, Junhao Shi, **Yang Xiao**, Mohan Jiang, Jie Sun, Yunze Wu, Shijie Xia, Xiaojie Cai, Tianze Xu, Weiye Si, Wenjie Li, Dequan Wang, Pengfei Liu.
   - *Under Review, 2026* (3rd Author).
   - [Code]

3. **LIMOPro: Reasoning Refinement for Efficient and Effective Test-time Scaling**

   - **Yang Xiao**, Jiashuo Wang, Ruifeng Yuan, Chunpu Xu, Kaishuai Xu, Wenjie Li, Pengfei Liu.
   - *NeurIPS 2025* (**First Author**).
   - Improved accuracy by up to 6.6% while reducing token usage by 3%-41% on AIME, AMC, GPQA benchmarks.
   - [Code]

4. **SCALE: Selective Resource Allocation for Overcoming Performance Bottlenecks in Mathematical Test-time Scaling**

   - **Yang Xiao**, Chunpu Xu, Ruifeng Yuan, Jiashuo Wang, Wenjie Li, Pengfei Liu.
   - *AAAI 2026* (**First Author**).
   - [Code]

5. **Limo: Less is More for Reasoning**

- Yixin Ye, Zhen Huang, **Yang Xiao**, Ethan Chern, Shijie Xia, Pengfei Liu.
- *COLM 2025* (3rd Author).
- Achieved 63.3% on AIME24 and 95.6% on MATH500 using only 1% of training data.
- [Code]

6. **Towards Dynamic Theory of Mind: Evaluating LLM Adaptation to Temporal Evolution of Human States**

- **Yang Xiao**, Jiashuo Wang, Qiancheng Xu, Changhe Song, Chunpu Xu, Yi Cheng, Wenjie Li, Pengfei Liu.
- *ACL 2025* (**First Author**).
- [Code]

7. **How Far Are We from Believable AI Agents? A Framework for Evaluating the Believability of Human Behavior Simulation**

- **Yang Xiao**, Yi Cheng, Jinlan Fu, Jiashuo Wang, Wenjie Li, Pengfei Liu.
- *arXiv Preprint* (**First Author**).

8. **Towards a Client-Centered Assessment of LLM Therapists by Client Simulation**

- Jiashuo Wang, **Yang Xiao**, Yanran Li, Changhe Song, Chunpu Xu, Chenhao Tan, Wenjie Li.
- *arXiv Preprint* (2nd Author).

SELECTED AWARDS AND HONORS

PolyU Presidential PhD Fellowship (Top 1%)	2023 – 2027
Outstanding Demo Paper Award at ACL	2022
Best Demo Award at ACL	2021
National Scholarship (Top 1%)	2018 – 2019