Yi Hu

huyi2002@stu.pku.edu.cn | Google Scholar | GitHub | Personal Website

Education

Peking University Sep 2024 - Jul 2029 (Expected)
PhD in Artificial Intelligence

Peking University Sep 2020 - Jul 2024
BSc in Physics

Research Interests

Personal Website

My research seeks to define the boundaries of large language models reasoning, revealing what can be reliably expected from them and what remains beyond their reach. I systematically investigate how LLMs reason, *where they fail*, *why failures occur*, and *how to fix them*. Specifically, my work focuses on:

- *Diagnosing systematic failure points* in LLM reasoning across mathematical, numerical, and physical domains;
- *Uncovering underlying mechanisms*, such as case-based / rule-based reasoning patterns and depth under-utilization;
- *Designing targeted improvements*, including training strategies that enhance strict rule-following and strengthen length generalization across varying tasks.

Publications

Google Scholar

Mechanistic Understanding of LLMs:

[1] Knowledge Interaction in Hypernetwork-based Continual Learning: Conflict and Composition *preprint 2026*
Jiaqi Zheng, **Yi Hu**, Yadong Mu, Muhan Zhang†

[2] Shared Weights, Selected Computations: How Looped Transformers Route What Each Loop Does *preprint 2026*
Jiaju Wu, **Yi Hu**, Muhan Zhang

[3] Towards a Mechanistic Understanding of Large Reasoning Models: A Survey of Training, Inference, and Failures *ACL 2026 Main (Oral)*
Yi Hu, Jiaqi Gu, Ruxin Wang, Zijun Yao, Hao Peng, Xiaobao Wu, Jianhui Chen, Muhan Zhang†, Liangming Pan†

[4] What Affects the Effective Depth of Large Language Models? *NeurIPS 2025 MI Workshop*
Yi Hu, Cai Zhou, Muhan Zhang†.

[5] Case-Based or Rule-Based: How Do Transformers Do the Math? *ICML 2024*
Yi Hu, Xiaojuan Tang, Haotong Yang, Muhan Zhang†.

Evaluating LLM Capabilities and Limitations:

[6] SubTokenTest: A Practical Benchmark for Real-World Sub-token Understanding *ACL 2026 Main*
Shuyang Hou*, **Yi Hu***, Muhan Zhang†

[7] Phybench: Holistic evaluation of physical perception and reasoning in large language models? *NeurIPS 2025*
Shi Qiu*, Shaoyang Guo*, Zhuo-Yang Song*, Yunbo Sun*, Zeyu Cai*, Jiashen Wei*, Tianyu Luo*, Yixuan Yin, Haoxu Zhang, **Yi Hu**, and others

[8] Number Cookbook: Number Understanding of Language Models and How to Improve It *ICLR 2025*

Haotong Yang, **Yi Hu**, Shijia Kang, Zhouchen Lin, Muhan Zhang†

Improving LLM Reasoning and Generalization:

[9] Coevolutionary continuous discrete diffusion: Make your diffusion language model a latent reasoner? *ICML 2026*

Cai Zhou, Chenxiao Yang, **Yi Hu**, Chenyu Wang, Chubin Zhang, Muhan Zhang, Lester Mackey, Tommi Jaakkola, Stephen Bates, Dinghuai Zhang†

[10] Beyond Single-Task: Robust Multi-Task Length Generalization for LLMs *NeurIPS 2025*

Yi Hu*, Shijia Kang*, Haotong Yang, Haotian Xu, Muhan Zhang†.

[11] Redstar: Does scaling long-cot data unlock better slow-reasoning systems? *preprint 2025*

Haotian Xu, Xing Wu, Weinong Wang, Zhongzhi Li, Da Zheng, Boyuan Chen, **Yi Hu**, and others

[12] Code prompting: a neural symbolic method for complex reasoning in large language models *preprint 2023*

Yi Hu, Haotong Yang, Zhouchen Lin, Muhan Zhang†

Experience

LLM Post-Training Team, Wizard Intelligence Learning Lab (WILL), Wizard Quant Beijing, China

Research Intern Mar 2026 – Now

- ***Research Focus:*** Primitive-level evaluation and targeted data synthesis for mathematical reasoning.
- ***Contributions:*** Built an agentic framework to extract and align mathematical primitives from problem corpora; developed primitive-level evaluation to track proficiency and identify capability gaps.
- Built a data-synthesis pipeline for underlearned primitives, generating problems with correctness and difficulty filtering and verified solutions.

LLM Post-Training Team, Xiaohongshu (RedNote) Beijing, China

Research Intern Oct 2024 – Jun 2025

- ***Research Focus:*** Improving logical reasoning and test-time scaling through post-training.
- ***Contributions:*** Collected and evaluated logical reasoning benchmarks, constructed long-chain reasoning datasets, and built an RL environment; experimented with SFT, DPO, and PPO to improve logical reasoning performance.
- Developed a specialized logical reasoning dataset and enhanced Qwen-32B, raising its logical reasoning performance to a level comparable with QwQ; contributed to the RedStar technical report.

LLM Research Group, Huawei Beijing, China

University-Industry Collaboration Project Dec 2024 – Dec 2025

- Integrated multi-turn RL from RAGEN into Align-Anything, a multimodal post-training framework; refactored environment interaction, reward computation, and policy update logic to work with its RL trainer.
- Adapted the multi-turn RL module to Huawei Ascend 910B NPUs and reproduced performance consistent with the original methodology.

Selected Honors and Awards

Outstanding Graduate Award, Peking University Jun 2024

Ubiquant Scholarship, Peking University Sep 2023

Xiaomi Scholarship, Peking University Sep 2022