

# Ruihan Yang

Handan Road 173, Hongkou District, Shanghai  
Postal Code: 200080

Email: [rhyang17@fudan.edu.cn](mailto:rhyang17@fudan.edu.cn)

Mobile: (+86)18915990288

Github: <https://github.com/rhyang2021>

## EDUCATION

---

**Ph.D. in Statistics** (Supervisor: Deqing Yang)

Fudan University, 2023-2027 (expected)

**Ph.D. in Applied Mathematics**

Fudan University, 2021-2023

**B.S. in Mathematics and Applied Mathematics**

Fudan University, 2017-2021

## EXPERIENCE

---

• **Moonshot AI RL Team**

Beijing, China

*Research Intern, Supervisor: Longhui Yu*

*Oct. 2025 - Present*

**Research Topic:** Developing agentic intelligence through reinforcement learning (K2.5 post-train)

• **Tencent Hunyuan AI Digital Human**

Shenzhen, China

*Research Intern, Supervisor: Fanghua Ye, Jian Li*

*Oct. 2024 - Oct. 2025*

**Research Topic:** Agent Learning

• **Tencent AI Lab**

Shenzhen, China

*Research Intern, Supervisor: Leyang Cui, Zhisong Zhang*

*March 2024 - Oct. 2024*

**Research Topic:** Long-Form Uncertainty Expression

## SELECTED AWARDS

---

• **National Scholarship for Graduate Students (Top 1%)**

*Oct. 2022*

• **Fudan University Excellent Student Award**

*2021-2025*

## PUBLICATIONS

---

[\*: Equal Contribution]

[1] **KIMI K2.5: VISUAL AGENTIC INTELLIGENCE**

Ruihan Yang as a contributor to Tool-Integrated Reasoning for Kimi K2.5 (**first HLE reach 50+**)

Technical Report, Preprint, 2025.

[2] **Think Fast and Slow: Step-Level Cognitive Depth Adaptation for LLM Agents**

Ruihan Yang, Fanghua Ye, Xiang Wei, Xinbo Xu, Xiaqiang Tang, Bo Zhao, Shanyi Wang, Zhaopeng Tu, Xiaolong Li, Deqing Yang

Long Paper, submitted to: ICML 2026.

[3] **Confidence Estimation for LLMs in Multi-turn Interactions**

Caiqi Zhang<sup>\*</sup>, Ruihan Yang<sup>\*</sup>, Xiaochen Zhu, Chengzu Li, Tiancheng Hu, Yijiang Dong, Deqing Yang, Nigel Collier

Long Paper, submitted to: ACL 2026.

[4] **ARIA: Training Language Agents with Intention-Driven Reward Aggregation**

Ruihan Yang<sup>\*</sup>, Yikai Zhang<sup>\*</sup>, Aili Chen, Siyu Yuan, Jiangjie Chen, Deqing Yang, Yanghua Xiao

Long Paper, in: NeurIPS 2025 (**Spotlight**).

[5] **The Lighthouse of Language: Enhancing LLM Agents via Critique-Guided Improvement**

Ruihan Yang, Fanghua Ye, Jian Li, Siyu Yuan, Yikai Zhang, Zhaopeng Tu, Xiaolong Li, Deqing Yang

Long Paper, in: NeurIPS 2025.

[6] **UNCLE: Benchmark Uncertainty in Long-form Generation**

Ruihan Yang<sup>\*</sup>, Caiqi Zhang<sup>\*</sup>, Zhisong Zhang, Xinting Huang, Nigel Collier, Dong Yu, Deqing Yang

Long Paper, in: EMNLP 2025.

[7] **LoGU: Long-form Generation with Uncertainty Expressions**

Ruihan Yang<sup>\*</sup>, Caiqi Zhang<sup>\*</sup>, Zhisong Zhang, Xinting Huang, Sen Yang, Nigel Collier, Dong Yu, Deqing Yang

Long Paper, in: ACL 2025.

[8] **SelfGoal: Your Language Agents Already Know How to Achieve High-level Goals**

Ruihan Yang, Jiangjie Chen, Yikai Zhang, Siyu Yuan, Aili Chen, Kyle Richardson, Yanghua Xiao, Deqing Yang

Long Paper, in: NAACL 2025.

[9] **Atomic Calibration of LLMs in Long-Form Generations**

Caiqi Zhang, Ruihan Yang, Zhisong Zhang, Xinting Huang, Sen Yang, Dong Yu, Nigel Collier

Long Paper, in: arXiv:2410.13246.

[10] **From Persona to Personalization: A Survey on Role-Playing Large Language Model Agents**  
Jiangjie Chen\*, Xintao Wang\*, Rui Xu\*, Siyu Yuan\*, Yikai Zhang\*, Wei Shi\*, Jian Xie\*, Shuang Li, Ruihan Yang, Tinghui Zhu, Aili Chen, Nianqi Li, Lida Chen, Caiyu Hu, Siye Wu, Scott Ren, Ziquan Fu, Yanghua Xiao  
Journal paper, in: TMLR 2024.

[11] **GumbelSoft: Diversified Language Model Watermarking via the GumbelMax-trick**  
Jiayi Fu, Xuandong Zhao, Ruihan Yang, Yuansen Zhang, Jiangjie Chen, Yanghua Xiao  
Long Paper, in: ACL 2024.

[12] **Computational and Neural Decomposition of Stop Signal Performance into Sustained Attention and Response Inhibition Elements: Theoretical and Clinical Implications**  
Yilin Zhang\*, Zhaoyu Zuo\*, Ruihan Yang, Chao Xie, Wenjie Hou, Shitong Xiang, Tianye Jia, Tobias Banaschewski, Robert Whelan, Jianfeng Feng, Gunter Schumann, Trevor Robbins, Qiang Luo  
Journal paper, Under Review, Brain.

[13] **Dynamic neural reconfiguration for distinct strategies during competitive social interactions**  
Ruihan Yang, Yina Ma, Bao-Bao Pan, Meghana A. Bhatt, Terry Lohrenz, Colin F. Camerer, Qiang Luo  
Journal paper, in: NeuroImage (IF=7.4).

## Academic Service

---

Reviewer: ACL 2024, 2025; EMNLP 2024, 2025