Hao Hu

mousehu1210@gmail.com mousehu.github.io/ Google Scholar

Working Experience

Moonshot AI *Oct. 2024 - Present*

Member of Technical Staff

- Lead the *Computer-Use Agent* project, overseeing the full lifecycle including environment development, RL framework design, data construction and model training (Mid-Training, SFT, Agentic RL). Achieved near human-level performance on the OSWorld benchmark (62.5% pass@1, 73.0% pass@4).
  - Architected *AgentGym*, a scalable, multi-modal, and low-latency environment for computer-use agents on VolcEngine, supporting over 50,000 concurrent sandboxes.
  - Developed *Pulse*, a high-throughput computer-use annotation platform, facilitating the collection of 40,000+ human-annotated trajectories for computer-use agents.
  - Established the training infrastructure and frameworks optimized for multi-turn and multi-modal agentic RL.
  - Constructed over 200B tokens of Mid-Training data for computer-use agents by filtering video data, open source datasets and self-generated trajectories, and curated over 20B tokens of high-quality SFT data from human annotations.
  - Contributed to *OpenRLVR*, an automated problem-generation system for computer-use agents, significantly enhancing the efficiency and diversity of training datasets.
- Enhanced *Kimi Deep Researcher* by significantly improving webpage generation quality and search capabilities.
- Contributed to *Kimi K1.5* and *Kimi K2* LLMs by delivering 348B high-quality math pre-training data recipes through rigorous document filtering, translation, and data scoring and classification pipelines.

Microsoft *Jun. 2018 - Sept. 2018*

Software Engineering Intern

Worked on newsletter recommendation systems at News & Relevance Team

- Automated cleaning of massive Bing news click logs to construct a training dataset for news recommendation system, raising a 20-class news classifier accuracy from 78.6% to 90.6% and meeting the production criteria.
- Contributed to the development and maintenance of *Sagitt*, the company-wide recommendation system visualization platform for real-time sample inspection and data analysis.

Education

Tsinghua University *Sept. 2019 - June. 2024*

Ph.D. in Computer Science, IIIS

Advisor: Prof. Chongjie Zhang and Prof. Yang Gao

Peking University *Sept. 2015 - July 2019*

B.S. in Computer Science, School of Electronics Engineering and Computer Science

B.S. in Theoretical and Applied Mechanics, School of Engineering

Publications

(* indicates equal contribution)

1. **OPENCUA: Open Foundations for Computer-Use Agents** [Paper] [Code]

   Xinyuan Wang, Bowen Wang*, Dunjie Lu*, Junlin Yang*, Tianbao Xie*,Junli Wang*,Jiaqi Deng, Xiaole Guo, Yiheng Xu,Chen Henry Wu, Zhennan Shen, Zhuokai Li, Ryan Li, Xiaochuan Li, Junda Chen,Boyuan Zheng, Peihang Liu, Fangyu Lei, Ruisheng Cao,Yeqiao Fu,Dongchan Shi,Martin Shi, Jiarui Hu,Yuyan Wang,Jixuan Chen,Yuxiao Ye,Danyang Zhang,**Hao Hu**,Huarong Chen, Dikang Du, Zaida Zhou, Haotian Yao, Ziwei Chen, Qizheng Gu, Yipu Wang, Heng Wang, Diyi Yang, Victor Zhong, Flood Sung, Y. Charles, Zhilin Yang, Tao Yu*

   Thirty-ninth Conference on Neural Information Processing Systems (**NeurIPS**), 2025

2. **Kimi K2: Open Agentic Intelligence** [Paper] [Blog]

   Kimi Team

   Technical Report, 2025

3. **Kimi K1.5: Scaling Reinforcement Learning with LLMs** [Paper]

   Kimi Team

   Technical Report, 2025

4. **Reason for Future, Act for Now: A Principled Architecture for Autonomous LLM Agents** [Paper] [Code] [Blog]

   Zhihan Liu*, **Hao Hu***, Shenao Zhang*, Hongyi Guo, Shuqi Ke, Boyi Liu, Zhaoran Wang

   Forty-first International Conference on Machine Learning (**ICML**), 2024

   NeurIPS Workshop on Foundation Models for Decision Making, 2023

5. **CLARIFY: Contrastive Preference Reinforcement Learning for Untangling Ambiguous Queries** [Paper] [Code]

   Ni Mu, **Hao Hu***, Xiao Hu, Yiqin Yang, Bo Xu, Qing-Shan Jia*

   Thirty-ninth Conference on Neural Information Processing Systems (**NeurIPS**), 2025

6. **Fewer May Be Better: Enhancing Offline Reinforcement Learning with Reduced Dataset** [Paper]

   *Yiqin Yang, Quanwei Wang, Chenghao Li, **Hao Hu**, Chengjie Wu, Yuhua Jiang, Dianyu Zhong, Ziyou Zhang, Qianchuan Zhao, Chongjie Zhang, Xu Bo*

   Forty-second International Conference on Machine Learning (**ICML**), 2025

7. **Episodic Novelty through Temporal Distance** [Paper] [Code]

   Hao Hu, Yiqin Yang*, Jianing Ye, Chengjie Wu, Ziqing Mai, Yujing Hu, Tangjie Lv, Changjie Fan, Qianchuan Zhao, Chongjie Zhang*

   Thirteenth International Conference on Learning Representations (**ICLR**), 2025

8. **Bayesian Design Principles for Offline-to-Online Reinforcement Learning** [Paper] [Code]

   Hao Hu*, Yiqin Yang*, Jianing Ye, Chengjie Wu, Ziqing Mai, Yujing Hu, Tangjie Lv, Changjie Fan, Qianchuan Zhao, Chongjie Zhang

   Forty-first International Conference on Machine Learning (**ICML**), 2024

9. **Planning, Fast and Slow: Online Reinforcement Learning with Action-Free Offline Data via Multiscale Planners** [Paper]

   Chengjie Wu*, **Hao Hu***, Yiqin Yang, Ning Zhang, Chongjie Zhang

   Forty-first International Conference on Machine Learning (**ICML**), 2024

10. **Stylized Offline Reinforcement Learning: Extracting Diverse High-Quality Behaviors from Heterogeneous Datasets** [Paper]

    Yihuan Mao, Chengjie Wu, Xi Chen, **Hao Hu**, Ji Jiang, Tianze Zhou, Tangjie Lv, Changjie Fan, Zhipeng Hu, Yi Wu, Yujing Hu, Chongjie Zhang

    Twelveth International Conference on Learning Representations (**ICLR**), 2024

11. **Unsupervised Behavior Extraction via Random Intent Priors** [Paper] [Code]

    Hao Hu*, Yiqin Yang*, Jianing Ye, Ziqing Mai, Chongjie Zhang

    Thirty-seventh Conference on Neural Information Processing Systems (**NeurIPS**), 2023

12. **One Objective to Rule Them All: A Maximization Objective Fusing Estimation and Planning for Exploration** [Paper] [Code]

    Zhihan Liu*, Miao Lu*, Wei Xiong*, Han Zhong, **Hao Hu**, Shenao Zhang, Sirui Zheng, Zhuoran Yang, Zhaoran Wang

    Thirty-seventh Conference on Neural Information Processing Systems (**NeurIPS**) **Spotlight**, 2023

13. **What is Essential for Unseen Goal Generalization of Offline Goal-conditioned RL?** [Paper] [Code]

    Ruiyang, Yong Lin, Xiaoteng Ma, **Hao Hu**, Chongjie Zhang, Tong Zhang

    Fortieth International Conference on Machine Learning (**ICML**), 2023

14. **The Provable Benefit of Unsupervised Data Sharing for Offline Reinforcement Learning** [Paper] [Code]

    Hao Hu*, Yiqin Yang*, Qianchuan Zhao, Chongjie Zhang

    Eleventh International Conference on Learning Representations (**ICLR**), 2023

15. **Flow to Control: Offline Reinforcement Learning with Lossless Primitive Discovery** [Paper] [Code]

    Yiqin Yang*, **Hao Hu***, Wenzhe Li*, Siyuan Li, Chongjie Zhang, Qianchuan Zhao

    Thirty-Seventh AAAI Conference on Artificial Intelligence (**AAAI**), 2023

16. **On the Role of Discount Factor in Offline Reinforcement Learning** [Paper] [Code]

    Hao Hu*, Yiqin Yang*, Qianchuan Zhao, Chongjie Zhang

    Thirty-ninth International Conference on Machine Learning (**ICML**), 2022

17. **Offline Reinforcement Learning with Value-based Episodic Memory** [Paper] [Code]

    Xiaoteng Ma*, Yiqin Yang*, **Hao Hu***, Qihan Liu, Jun Yang, Chongjie Zhang, Qianchuan Zhao, Bin Liang

    Tenth International Conference on Learning Representations (**ICLR**), 2022

18. **On the Estimation Bias in Double Q-Learning** [Paper] [Code]

    Zhizhou Ren, Guangxiang Zhu, **Hao Hu**, Beining Han, Jianglun Chen, Chongjie Zhang

Thirty-fifth Conference on Neural Information Processing Systems (**NeurIPS**), 2021

19. **MetaCURE: Meta Reinforcement Learning with Empowerment-Driven Exploration** [Paper] [Code]

    Jin Zhang*, Jianhao Wang*, **Hao Hu**, Tong Chen, Yingfeng Chen, Changjie Fan, Chongjie Zhang

    Thirty-eighth International Conference on Machine Learning (**ICML**), 2021

20. **Generalizable Episodic Memory for Deep Reinforcement Learning** [Paper] [Code]

    Hao Hu, Jianing Ye, Zhizhou Ren, Guangxiang Zhu, Chongjie Zhang

    Thirty-eighth International Conference on Machine Learning (**ICML**), 2021

21. **Query-Efficient Offline Preference-Based Reinforcement Learning via In-Dataset Exploration** [Paper]

    Hao Hu*, Yiqin Yang*, Shuai Wang, Bo Liu, Yang Gao, Chongjie Zhang

    Under Review

Research Experience

MIG Group, Washington University in St. Louis *June. 2024 - Oct. 2024*
Worked on offline reinforcement learning
Advisor: Prof. Chongjie Zhang

NURL Group, IEMS, Northwestern University *Apr. 2023 - Sept. 2023*
Worked on reinforcement learning theory and building agentic systems with large language models
Advisor: Prof. Zhaoran Wang

MIG Group, IIIS, Tsinghua University *Sept. 2019 - Present*
Worked on episodic memory, offline reinforcement learning, semi-supervised reinforcement learning
Advisor: Prof. Chongjie Zhang and Prof. Yang Gao

Vision and Media Computing Group, Peking University *Nov. 2017 - Jun. 2018*
Worked on semantic segmentation and autonomous driving
Advisor: Prof. Shiliang Zhang

Selected Talks

- Data-Driven Reinforcement Learning [Slides] *Nov. 2023*
  Bytedance AI Lab
- Unsupervised Behavior Extraction via Random Intent Priors [Slides] *Oct. 2023*
  The 64th Seminar of Reinforcement Learning China Community (RL China)
- On the Role of Discount Factor in Offline Reinforcement Learning [Slides] [Video] *Jun. 2022*
  The 21st Seminar of Reinforcement Learning China Community (RL China)
- Generalizable Episodic Memory for Deep Reinforcement Learning [Slides] *Jul. 2021*
  Nanjing University