Zhihong Shao

E-mail: zhshaothu@gmail.com
Phone: +86 13121259158
Web: https://ZhihongShao.github.io

Research Interests

My interests are in natural language processing and deep learning. I am particularly interested in how we can build a robust and scalable AI system that can leverage diverse skills (e.g., tool use and reasoning) to aggregate possibly-heterogeneous information and answer natural language questions precisely regardless of their complexity.

Education

Tsinghua University, Beijing, China — September 2019 - July 2024
Ph.D. Student, Computer Science and Technology
Advisor: Minlie Huang

Beihang University, Beijing, China — September 2015 July 2019
B.E., Computer Science and Technology
GPA: 3.86/4, *Rank:* 2/213

Research Highlights

LLM Reasoning & Tool Augmentation

- `Informal Math Pre-Training and (large-scale) RL`: The DeepSeekMath [19] project demonstrates an effective data engineering pipeline for math pre-training, and lays the GRPO-based RL foundation for post-training DeepSeek models. The DeepSeek-R1 [15] project further leverages large-scale RL to create a strong reasoning model that approaches OpenAI's o1 performance in many reasoning tasks;
- `Formal Math Data Synthesis and Proof Search`: The DeepSeek-Prover project improves formal math reasoning (i.e., to generate math proofs that can be automatically verified) with large-scale expert iteration [2], RL from proof assistant's feedback [1][14], and tree search [1];
- `Reasoning with Tool Integration`: The ToRA [5] project augments chain-of-thought reasoning with Python code for strong math performance. The CRITIC [7] project experiments on more general reasoning tasks to study self-correction based on feedback from tools.

Publications

[1] DeepSeek-Prover-V1.5: Harnessing Proof Assistant Feedback for Reinforcement Learning and Monte-Carlo Tree Search
Huajian Xin*, Z.Z. Ren*, Junxiao Song*, **Zhihong Shao***, DeepSeek-AI
International Conference on Learning Representations (ICLR), 2025.

[2] DeepSeek-Prover: Advancing Theorem Proving in LLMs through Large-Scale Synthetic Data
Huajian Xin, Daya Guo, **Zhihong Shao**, Zhizhou Ren, Qihao Zhu, Bo Liu, Chong Ruan, Wenda Li, Xiaodan Liang
Annual Conference on Neural Information Processing Systems (NeurIPS), MATH-AI workshop, 2024.

[3] Math-Shepherd: Verify and Reinforce LLMs Step-by-step without Human Annotations
Peiyi Wang, Lei Li, **Zhihong Shao**, R.X. Xu, Damai Dai, Yifei Li, Deli Chen, Y.Wu, Zhifang Sui
Annual Meeting of the Association for Computational Linguistics (ACL), 2024.

[4] Learning Task Decomposition to Assist Humans in Competitive Programming
Jiaxin Wen, Ruiqi Zhong, Pei Ke, **Zhihong Shao**, Hongning Wang, Minlie Huang
Annual Meeting of the Association for Computational Linguistics (ACL), 2024.

[5] ToRA: A Tool-Integrated Reasoning Agent for Mathematical Problem Solving
Zhihong Shao*, Zhibin Gou*, Yeyun Gong, Yelong Shen, Yujiu Yang, Minlie Huang, Nan Duan, Weizhu Chen
International Conference on Learning Representations (ICLR), 2024.

[6] Enhancing Retrieval-Augmented Large Language Models with Iterative Retrieval-Generation Synergy
Zhihong Shao, Yeyun Gong, Yelong Shen, Minlie Huang, Nan Duan, Weizhu Chen
Findings of Empirical Methods in Natural Language Processing (Findings of EMNLP), 2023.

[7] CRITIC: Large Language Models Can Self-Correct with Tool-Interactive Critiquing
Zhibin Gou, **Zhihong Shao**, Yeyun Gong, Yelong Shen, Yujiu Yang, Nan Duan, Weizhu Chen
International Conference on Learning Representations (ICLR), 2024.

[8] Synthetic Prompting: Generating Chain-of-Thought Demonstrations for Large Language Models
Zhihong Shao, Yeyun Gong, Yelong Shen, Minlie Huang, Nan Duan, and Weizhu Chen
International Conference on Machine Learning (ICML), 2023.

[9] Chaining Simultaneous Thoughts for Numerical Reasoning
Zhihong Shao, Fei Huang, and Minlie Huang
Findings of Empirical Methods in Natural Language Processing (Findings of EMNLP), 2022.

[10] Answering Open-Domain Multi-Answer Questions via a Recall-then-Verify Framework
Zhihong Shao, and Minlie Huang
Annual Meeting of the Association for Computational Linguistics (ACL), 2022.

[11] AdvExpander: Generating Natural Language Adversarial Examples by Expanding Text
Zhihong Shao, Zhongqin Wu, and Minlie Huang
IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), vol. 30, pp. 1184-1196, 2022.

[12] A Mutual Information Maximization Approach for the Spurious Solution Problem in Weakly Supervised Question Answering
Zhihong Shao, Lifeng Shang, Qun Liu, and Minlie Huang
Annual Meeting of the Association for Computational Linguistics (ACL), 2021.

[13] Long and Diverse Text Generation with Planning-based Hierarchical Variational Model
Zhihong Shao, Minlie Huang, Jiangtao Wen, Wenfei Xu, and Xiaoyan Zhu
Empirical Methods in Natural Language Processing (EMNLP), 2019.

Preprint

[14] DeepSeek-Prover-V2: Advancing Formal Mathematical Reasoning via Reinforcement Learning for Subgoal Decomposition
Z.Z. Ren*, **Zhihong Shao***, Junxiao Song*, DeepSeek-AI
Arxiv abs/2504.21801, 2025.

[15] DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning
DeepSeek-AI
Arxiv abs/2501.12948, 2025.

[16] DeepSeek-V3 Technical Report
DeepSeek-AI
Arxiv abs/2412.19437, 2024.

[17] DeepSeek-Coder-V2: Breaking the Barrier of Closed-Source Models in Code Intelligence
Qihao Zhu*, Daya Guo*, **Zhihong Shao***, Dejian Yang*, DeepSeek-AI
Arxiv abs/2406.11931, 2024.

[18] DeepSeek-V2: A Strong, Economical, and Efficient Mixture-of-Experts Language Model
DeepSeek-AI
Arxiv abs/2405.04434, 2024.

[19] DeepSeekMath: Pushing the Limits of Mathematical Reasoning in Open Language Models
Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Mingchuan Zhang, Y.K. Li, Y. Wu, Daya Guo
Arxiv abs/2402.03300, 2024.

[20] DeepSeek LLM: Scaling Open-Source Language Models with Longtermism
DeepSeek-AI
Arxiv abs/2401.02954, 2024.

[21] CoTK: An Open-Source Toolkit for Fast Development and Fair Evaluation of Text Generation
Fei Huang, Dazhen Wan, **Zhihong Shao**, Pei Ke, Jian Guan, Yilin Niu, Xiaoyan Zhu, and Minlie Huang
Arxiv abs/2002.00583, 2020.

Awards

Lenovo Scholarship, Tsinghua University 2023
1st Prize, Comprehensive Scholarship, Tsinghua University 2022
2nd Prize, Comprehensive Scholarship, Tsinghua University 2021
3rd Prize, the National Final of "LAN QIAO CUP" C/C++ Group 2018
1st Prize, National College Students Mathematics Competition (non-math-major) 2016
China National Scholarship 2016, 2017, 2018

Services

Reviewer/Program Committee: ACL, EMNLP, NLPCC, ARR

Teaching Assistant

Artificial Neural Network Fall 2019 - 2022
Instructor: Minlie Huang

Object-Oriented Programming Spring 2020 - 2023
Instructor: Minlie Huang
Also gave guest lectures and made assignments