
ZERU SHI

CONTACT INFORMATION

Ph.D. Student, Department of Computer Science, Rutgers University

HOME PAGE: Zeru_Shi homepage

PHONE: (+1)(732)354-2576 EMAIL: zeru.shi@rutgers.edu

RESEARCH INTERESTS

- **AI interpretability:** I aim to use interpretability analyses to uncover underlying phenomena in LLMs and, based on these insights, design simple yet effective methods to enhance their reasoning ability. I am also interested in interpretable reinforcement learning for LLMs, seeking to improve post-training performance through a better understanding of model behavior.
- **LLM/VLM reasoning:** I am highly interested in reasoning within LLMs and VLMs. My work explores post-training methods, such as reinforcement learning based optimization, as well as training-free approaches to improve their reasoning capabilities.
- **LLM Agent:** I am interested in LLM agents particularly in the area of agent memory. I have worked on using vector databases to efficiently manage memory in multi-agent systems.

EDUCATION

Rutgers University, New Brunswick, USA

September 2025 - May 2030(Expected)

Ph.D. in Computer Science

Department of Computer Science

Advisor: Dr. Ruixiang Tang

Dalian University of Technology, Liaoning, China

September 2021 - June 2025

Bachelor of Science in Software Engineering

Department of Software Engineering

Advisor: Dr. Risheng Liu

SELECTED HONORS AND AWARDS

- Dalian excellent undergraduate graduates, China
- National Scholarship, China
- Special scholarship of NOK Corporation, Japan

PUBLICATIONS

* denotes an equal contribution.

[ICLR' 25] **Zeru Shi***, Kai Mei*, Mingyu Jin, Yongye Su, Chaoji Zuo, et al. "From Commands to Prompts: LLM-based Semantic File System" In *The International Conference on Learning Representations*

- This work focuses on **agent memory**. It introduces a *vector database* based data management mechanism for multi-agent systems, enabling users to store local files into an agent's memory through natural language instructions and manage that memory entirely within the agent system, without requiring Linux command line operations. Compared to not using our system, the average accuracy of file retrieval improves by **41%**, while the average time cost is reduced by **80%**. The framework also provides an open interface, allowing it to be easily integrated into different systems.

[Under Review] **Zeru Shi**, Yingjia Wan, Zhenting Wang, Qifan Wang, Fan Yang, Elisa Kreiss, Ruixiang Tang. "Meaningless Tokens, Meaningful Gains: How Activation Shifts Enhance LLM Reasoning". In submission.

- This work focuses on **LLM reasoning and XAI**. It is motivated by the phenomenon known as the *meaningless token effect*, which observes that inserting long sequences of repeated meaningless tokens into the input can unexpectedly improve, rather than degrade, model performance. Through detailed analysis, we find that such repeated meaningless tokens induce an *activation redistribution in the MLP layers: activations near zero become less frequent, while larger activations occur more often for meaningful tokens*. Building on this insight, we distill the phenomenon into a simpler, more efficient method that improves reasoning performance by an average of **2.7** without increasing inference time. We further validate the effectiveness of this approach across multiple datasets and model architectures.

[IEEE TCSVT] **Zeru Shi**, Zengxi Zhang, Kemeng Cui, Ruizhe An, Jinyuan Liu, Zhiying Jiang. “Se-FENet: Robust Deep Homography Estimation via Semantic-Driven Feature Enhancement” In *IEEE Transactions on Circuits and Systems for Video Technology (IF = 11.1)*

- This work relates to low-level computer vision. It introduces meta learning techniques into neural networks to enhance the model’s homography estimation capability under challenging and adverse environments.

[EMNLP’ 25] Yongye Su, Yucheng Zhang, **Zeru Shi**, Bruno Ribeiro, Elisa Bertino “Castle: Causal Cascade Updates in Relational Databases with Large Language Models” In *Empirical Methods in Natural Language Processing (Main)*

- This work focuses on **applying LLMs to relational databases**. It proposes a *cascaded-update mechanism that enables natural-language instructions to trigger multi-column, causally consistent SQL UPDATE statements without exposing table contents to the model*. The system determines not only which columns require direct modification but also how these updates propagate through the schema, inducing cascading updates via nested queries and structured dependencies, all while preserving data confidentiality. For casual cascade, the **F-1 score improve at least 21%**.

[Under Review] **Zeru Shi**, Zhenting Wang, Yongye Su, Weidi Luo, Hang Gao, et al. “Auto-Prompt Generation is Not Robust: Prompt Optimization Driven by Pseudo Gradient”. In submission.

- This work focuses on **prompt optimization** under input perturbations, *improving model performance when user inputs contain typos or imprecise expressions*. It also introduces a corresponding dataset that includes various forms of erroneous or underspecified inputs. Finally, the method is validated across multiple models on tasks such as text summarization, simplification, and classification, demonstrating its effectiveness. Especially for summarization tasks, where performance improves by **at least 10%**.

INTERNSHIP EXPERIENCE

Research Scientist Intern, Shanghai AI Laboratory
AI for Science Group

November 2024 - April 2025

- During my internship, I worked on **generative model**, with a focus on 3D generation and perception. I contributed to the development and training of a large scale any-to-3D model, participating in both system design and model training workflows.

MISCS

Programming Skills ans Tools: Python(Strong), C/C++(Proficient), MySQL(Proficient)

Language: English, Chinese, Japanese