Curriculum Vitae

Personal Information

Ruisheng Cao
June 23rd, 1996
rhythmcao
ruishengcao@gmail.com
+86 18217583036
rhythmcao.github.io

Education

- PhD candidate in Computer Technology, Shanghai Jiao Tong University *2021-present*
- The 3rd Wu Wenjun AI Honorary Doctoral Program (only **15** students in SJTU) *2021-present*
- Intern at XLANG Lab, the University of Hong Kong *Jan-Apr 2024*
- M.Eng. in Computer Technology, Shanghai Jiao Tong University *2018-2021*
  - GPA : 3.94/4.0 Rankings : 3
- Zhiyuan Honors Program of Engineering (an elite program for **top 5**% students) *2014-2018*
- Summer Program at Department of Computer Science, Cornell University *Jun-Jul 2017*
- Humanities Program at Hertford College part of the University of Oxford *Aug 2015*
- B.Eng. in Computer Science and Technology, Shanghai Jiao Tong University *2014-2018*
  - GPA : 3.87/4.3 Rankings : top 10%

Honors and Awards

Wu Wenjun AI Honorary Doctoral Scholarship *2022-2024*

Outstanding Graduate of Shanghai Jiao Tong University *2021*

National Scholarship for Graduate Students *2018-2019, 2019-2020*

Zhiyuan Honor Degree of Bachelor of Science (only **48** undergraduates in SJTU) *2018*

Zhiyuan Honors Scholarship *2014-2015, 2015-2016*

Academic Excellence Scholarship (First-Class, top 4) *2014-2015*

Research Interests

Research interests mainly lie in structured natural language understanding, model-based data generation with iterative training, and LLM or VLM-based multi-modal agents. The already published work focuses on 1) dedicated model architecture which integrates structured knowledge into neural-symbolic processing tasks such as text-to-SQL, semantic parsing and task-oriented dialogue system, 2) pseudo data augmentation which leverages the dual model with cycle learning, and 3) GUI interface interaction in fields like coding, data science and engineering, and computer control.

Selected Publications

Currently, 5 conference papers and 2 journal articles as first or co-first author. See for full list.

- Fangyu Lei, Jixuan Chen, Yuxiao Ye, **Ruisheng Cao**, Dongchan Shin, Hongjin SU, ZHAOQING SUO, Hongcheng Gao, Wenjing Hu, Pengcheng Yin, Victor Zhong, Caiming Xiong, Ruoxi Sun, Qian Liu, Sida Wang, and Tao Yu. *Spider 2.0 : Can Language Models Resolve Real-World Enterprise Text-to-SQL Workflows ?* (**Top 10 Highest-Scoring Papers**) Submitted to International Conference on Learning Representations (ICLR), 2025.
- **Ruisheng Cao**, Fangyu Lei, Haoyuan Wu, Jixuan Chen, Yeqiao Fu, Hongcheng Gao, Xinzhuang Xiong, Hanchong Zhang, Yuchen Mao, Wenjing Hu, Tianbao Xie, Hongshen Xu, Danyang Zhang, Sida Wang, Ruoxi Sun, Pengcheng Yin, Caiming Xiong, Ansong Ni, Qian Liu, Victor Zhong, Lu Chen, Kai Yu and Tao Yu. *Spider2-V : How Far Are Multimodal Agents From Automating Data Science and Engineering Workflows ?* (**Spotlight**) Advances in Neural Information Processing Systems (NeurIPS), 2024.
- Hanchong Zhang, **Ruisheng Cao**, Hongshen Xu, Lu Chen and Kai Yu. *CoE-SQL : In-Context Learning for Multi-Turn Text-to-SQL with Chain-of-Editions.* In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies (NAACL), pages 6487-6508, Mexico City, Mexico.
- Tianbao Xie, Danyang Zhang, Jixuan Chen, Xiaochuan Li, Siheng Zhao, **Ruisheng Cao**, Toh Jing Hua, Zhoujun Cheng, Dongchan Shin, Fangyu Lei, Yitao Liu, Yiheng Xu, Shuyan Zhou, Silvio Savarese, Caiming Xiong, Victor Zhong and Tao Yu. *OSWorld : Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments.* Advances in Neural Information Processing Systems (NeurIPS), 2024.

Selected Publications

- **Ruisheng Cao**, Lu Chen, Jieyu Li, Hanchong Zhang, Hongshen Xu, Wangyou Zhang and Kai Yu. *A Heterogeneous Graph to Abstract Syntax Tree Framework for Text-to-SQL*. In IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2023 (01) : 1-16.
- Hanchong Zhang, **Ruisheng Cao**, Lu Chen, Hongshen Xu and Kai Yu. *ACT-SQL : In-Context Learning for Text-to-SQL with Automatically-Generated Chain-of-Thought.* In Findings of the Association for Computational Linguistics : EMNLP 2023, pages 3501-3532, Singapore. Association for Computational Linguistics.
- **Ruisheng Cao*** (equal contribution, responsible for code), Hongshen Xu*, Su Zhu, Sheng Jiang, Hanchong Zhang, Lu Chen, Kai Yu. *A Birgat Model for Multi-Intent Spoken Language Understanding with Hierarchical Semantic Frames.* ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2024 : 12251-12255.
- Hanchong Zhang, Jieyu Li, Lu Chen, **Ruisheng Cao**, Yunyan Zhang, Yu Huang, Yefeng Zheng, and Kai Yu. *CSS : A Large-scale Cross-schema Chinese Text-to-SQL Medical Dataset.* In Findings of the Association for Computational Linguistics : ACL 2023, pages 6970-6983, Toronto, Canada. Association for Computational Linguistics.
- **Ruisheng Cao**, Lu Chen, Zhi Chen, Yanbin Zhao, Su Zhu and Kai Yu. *LGESQL : Line Graph Enhanced Text-to-SQL Model with Mixed Local and Non-Local Relations* (**Long Oral**). In 59th Annual Meeting of the Association for Computational Linguistics (ACL) and the 11th International Joint Conference on Natural Language Processing, pages 2541-2555, 2021.
- Zihan Zhao, Lu Chen, **Ruisheng Cao**, Hongshen Xu, Xingyu Chen and Kai Yu. *TIE : Topological Information Enhanced Structural Reading Comprehension on Web Pages.* Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies (NAACL). Seattle, United States, pages 1808-1821, 2022.
- Su Zhu, Lu Chen, **Ruisheng Cao**, Zhi Chen, Qingliang Miao and Kai Yu. *Few-Shot NLU with Vector Projection Distance and Abstract Triangular CRF.* Natural Language Processing and Chinese Computing : 10th CCF International Conference, NLPCC 2021, Qingdao, China, October 13-17, 2021, Proceedings, Part I 10. Springer International Publishing, 2021 : 505-516.
- Zhi Chen, Lu Chen, Hanqi Li, **Ruisheng Cao**, Da Ma, Mengyue Wu and Kai Yu. *Decoupled Dialogue Modeling and Semantic Parsing for Multi-Turn Text-to-SQL.* Findings of the Association for Computational Linguistics : ACL-IJCNLP, pages 3063-3074, 2021.
- Zhi Chen, Lu Chen, Yanbin Zhao, **Ruisheng Cao**, Zihan Xu, Su Zhu and Kai Yu. *ShadowGNN : Graph Projection Neural Network for Text-to-SQL Parser.* Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies (NAACL), online, pages 5567-5577, 2021.
- **Ruisheng Cao**, Su Zhu, Chenyu Yang, Chen Liu, Rao Ma, Yanbin Zhao, Lu Chen and Kai Yu. *Unsupervised Dual Paraphrasing for Two-stage Semantic Parsing* (**Long Online Oral**). In 58th Annual Meeting of the Association for Computational Linguistics (ACL), Seattle, USA, pages 6806-6817, 2020.
- **Ruisheng Cao*** (equal contribution, responsible for code), Su Zhu* and Kai Yu. *Dual Learning for Semi-Supervised Natural Language Understanding.* IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1936-1947, 2020.
- Yanbin Zhao, Lu Chen, Zhi Chen, **Ruisheng Cao**, Su Zhu and Kai Yu. *Line Graph Enhanced AMR-to-Text Generation with Mix-Order Graph Attention Networks* (**Long Online Oral**). In 58th Annual Meeting of the Association for Computational Linguistics (ACL), Seattle, USA, pages 732-741, 2020.
- **Ruisheng Cao**, Su Zhu, Chen Liu, Jieyu Li and Kai Yu. *Semantic Parsing with Dual Learning* (**Long Oral**). In 57th Annual Meeting of the Association for Computational Linguistics (ACL), Florence, Italy, pages 51-64, 2019.

Challenges And Competitions

Rank 1 DuSQL : A Large-Scale and Pragmatic Chinese Text-to-SQL Dataset. *2021/10-present*

Rank 1 Spider : Yale Semantic Parsing and Text-to-SQL Challenge. *2021/06-2022/09*

Rank 1 CSpider : The Chinese Semantic Parsing and Text-to-SQL Challenge. *2022/05-2023/06*

Rank 1 NL2SQL : A Large-Scale Chinese Single-Table Text-to-SQL Dataset. *2022/04-present*

Rank 2 The 9th China National Conference on Social Media Processing (SMP 2020). The Evaluation of Chinese Human-Computer Dialogue Technology, ECDT : Few Shot Spoken Language Understanding. *2020*