

Ziru “Ron” Chen

Curriculum Vitae

Last Updated: October 18, 2025

Email: chen.8336@osu.edu

Website: <https://ronch99.github.io/>

Twitter: @RonZiruChen

INTERESTS

I am broadly interested in *conversational AI agents*, *natural language processing*, and *machine learning*. My research focuses on (1) improving **post-training and test-time computation** of language models and (2) building **conversational agents for coding and data analysis**, with applications to complex real-world tasks such as database management and scientific R&D.

EDUCATION

The Ohio State University, Columbus, OH Aug 2022 – May 2027
Ph.D. in Computer Science and Engineering **GPA: 3.92 / 4.0**

The Ohio State University, Columbus, OH Aug 2018 – May 2022
B.S. in Computer and Information Science, Math Minor **GPA: 3.97 / 4.0**
Summa Cum Laude

EXPERIENCES

Research Assistant, The Ohio State University Dec 2020 – Present
Advisor: Dr. Huan Sun

- Test-Time Computation of Language Models (**ACL’24**)
- Language Agents for Coding and Scientific Data Analysis (**ICLR’25**, **EMNLP’25**)
- Interactive Text-to-SQL Generation (**ACL’23**, **EMNLP’23**, **EMNLP Findings’23**)
- User-Centered Task-Oriented Dialogue System (**Alexa Prize’22**, **3rd Place Honor**)

Research Intern, Microsoft GenAI May 2025 – Aug 2025
Mentors: Dr. Yujia Xie, Dr. Dongdong Chen

- Reinforcement Learning for Multi-Turn Code Reasoning and Generation

HONORS

University Fellowship, The Ohio State University 2022–2023, 2026–2027
3rd Place (Student Co-lead), 1st Alexa Prize Taskbot Challenge, Amazon 2022
Undergraduate Research Award, The Ohio State University 2022
CIS Undergrad Scholarship, The Ohio State University 2021

PUBLICATIONS

[17] *Holistic Agent Leaderboard: The Missing Infrastructure for AI Agent Evaluation*
Sayash Kapoor, Benedikt Stroebl, Peter Kirgis, Nitya Nadgir, Zachary S Siegel, Boyi Wei, Tianci Xue, **Ziru Chen**, Felix Chen, Saiteja Utpala, Franck Ndzomga, Dheeraj Oruganty, Sophie Luskin, Kangheng Liu, Botao Yu, Amit Arora, Dongyoon Hahm, Harsh Trivedi, Huan Sun, Juyong Lee, Tengjun Jin, Yifan Mai, Yifei Zhou, Yuxuan Zhu, Rishi Bommasani, Daniel Kang, Dawn Song, Peter Henderson, Yu Su, Percy Liang, Arvind Narayanan
In Submission to ICLR 2026

[16] *Large Language Models Achieve Gold Medal Performance at International Astronomy & Astrophysics Olympiad*
Lucas Carrit Delgado Pinheiro*, **Ziru Chen***, Bruno Caixeta Piazza, Ness Shroff, Yingbin Liang, Yuan-Sen Ting, Huan Sun (* equal contribution)
In Submission to Nature Astronomy

[15] *Agent Data Protocol: Unifying Datasets for Diverse, Effective Fine-tuning of LLM Agents*
Yueqi Song, Ketan Ramaneti, Zaid Sheikh, **Ziru Chen**, Boyu Gou, Tianbao Xie, Yiheng Xu, Danyang Zhang, Apurva Gandhi, Fan Yang, Joseph Liu, Tianyue Ou, Zhihao Yuan, Frank F. Xu, Shuyan Zhou, Xingyao Wang, Xiang Yue, Tao Yu, Huan Sun, Yu Su, Graham Neubig
In Submission to ICLR 2026

[14] *AutoSDT: Scaling Data-Driven Discovery Tasks Toward Open Co-Scientists*
Yifei Li*, Hanane Nour Moussa*, **Ziru Chen**, Shijie Chen, Botao Yu, Mingyi Xue, Benjamin Burns, Tzu-Yao Chiu, Vishal Dey, Zitong Lu, Chen Wei, Qianheng Zhang, Tianyu Zhang, Song Gao, Xuhui Huang, Xia Ning, Nesreen K. Ahmed, Ali Payani, Huan Sun (* equal contribution)
The 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP 2025, Long)
Best Paper Award, The first LLM for Scientific Discovery Workshop at COLM 2025

- [13] *GeoAnalystBench: A GeoAI benchmark for assessing large language models for spatial analysis workflow and code generation*
 Qianheng Zhang, Song Gao, Chen Wei, Yibo Zhao, Ying Nie, **Ziru Chen**, Shijie Chen, Yu Su, Huan Sun
 Transactions in GIS 2025 (TGIS 2025, Journal)
- [12] *ScienceAgentBench: Toward Rigorous Assessment of Language Agents for Data-Driven Scientific Discovery*
Ziru Chen, Shijie Chen, Yuting Ning, Qianheng Zhang, Boshi Wang, Botao Yu, Yifei Li, Zeyi Liao, Chen Wei, Zitong Lu, Vishal Dey, Mingyi Xue, Frazier N. Baker, Benjamin Burns, Daniel Adu-Ampratwum, Xuhui Huang, Xia Ning, Song Gao, Yu Su, Huan Sun
 The Thirteenth International Conference on Learning Representations (ICLR 2025, Long)
Covered by Nature News. 9.1K+ downloads on Huggingface by October 18, 2025.
- [11] *Tooling or Not Tooling? The Impact of Tools on Language Agents for Chemistry Problem Solving*
 Botao Yu, Frazier N. Baker, **Ziru Chen**, Garrett Herb, Boyu Gou, Daniel Adu-Ampratwum, Xia Ning, Huan Sun
 Findings of the Association for Computational Linguistics: NAACL 2025 (Findings 2025, Short)
- [10] *When is Tree Search Useful for LLM Planning? It Depends on the Discriminator*
Ziru Chen, Michael White, Raymond Mooney, Ali Payani, Yu Su, Huan Sun
 The 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024, Long)
- [9] *eCeLLM: Generalizing Large Language Models for E-commerce from Large-scale, High-quality Instruction Data*
 Bo Peng*, Xinyi Ling*, **Ziru Chen**, Huan Sun, Xia Ning (* equal contribution)
 Proceedings of the 41st International Conference on Machine Learning (ICML 2024, Long)
Our ECInstruct dataset is recommended at Amazon KDD Cup 2024.
- [8] *Automatic Evaluation of Attribution by Large Language Models*
 Xiang Yue, Boshi Wang, **Ziru Chen**, Kai Zhang, Yu Su, Huan Sun
 Findings of the Association for Computational Linguistics: EMNLP 2023 (Findings 2023, Long)
- [7] *Error Detection for Text-to-SQL Semantic Parsing*
 Shijie Chen, **Ziru Chen**, Huan Sun, Yu Su
 Findings of the Association for Computational Linguistics: EMNLP 2023 (Findings 2023, Long)
- [6] *Exploring Chain-of-Thought Style Prompting for Text-to-SQL*
 Chang-Yu Tai*, **Ziru Chen***, Tianshu Zhang, Xiang Deng, Huan Sun (* equal contribution)
 The 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023, Long)
- [5] *Roll Up Your Sleeves: Working with a Collaborative and Engaging Task-Oriented Dialogue System*
 Lingbo Mo, Shijie Chen, **Ziru Chen**, Xiang Deng, Ashley Lewis, Sunit Singh, Samuel Stevens, Chang-Yu Tai, Zhen Wang, Xiang Yue, Tianshu Zhang, Yu Su, Huan Sun
 The 24th Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL 2023, Demo)
- [4] *Text-to-SQL Error Correction with Language Models of Code*
Ziru Chen, Shijie Chen, Michael White, Raymond Mooney, Ali Payani, Jayanth Srinivasa, Yu Su, Huan Sun
 The 61st Annual Meeting of the Association for Computational Linguistics (ACL 2023, Short)
- [3] *SalsaBot: Towards a Robust and Generalizable Embodied Agent*
 Chan Hee Song*, Jiaman Wu*, Ju-Seung Byun, Zexin Xu, Vardaan Pahuja, Goonmeet Bajaj, Samuel Stevens, **Ziru Chen**, Yu Su (* equal contribution)
 The 1st Proceedings of Alexa Prize Simbot Challenge (Alexa Prize 2023, Long)
 The Embodied AI 2023 Workshop (EmbodiedAI 2023, Short)
- [2] *Error Detection for Interactive Text-to-SQL Semantic Parsing*
 Shijie Chen, **Ziru Chen**, Huan Sun, Yu Su

[1] *Bootstrapping a User-Centered Task-Oriented Dialogue System*

Shijie Chen*, **Ziru Chen***, Xiang Deng, Ashley Lewis, Lingbo Mo, Samuel Stevens, Zhen Wang, Xiang Yue, Tianshu Zhang, Yu Su, Huan Sun (* equal contribution)

The 1st Proceedings of Alexa Prize Taskbot Challenge (Alexa Prize 2022, Long)

Our TacoBot is the third-place winner and deployed in Amazon Alexa for one year.

TEACHING

CSE 5525 Speech and Language Processing, Undergraduate & Graduate Level

Teaching Assistant, The Ohio State University (Autumn 2021; Spring 2022)

CSE 2321 Foundations I: Discrete Structures, Undergraduate Level

Teaching Assistant, The Ohio State University (Autumn 2019; Spring, Autumn 2020)

SERVICE

Outstanding Reviewer: EMNLP 2024 (ACL Rolling Review 06/2024)

Program Committee: AAAI (2023, 2024, 2025), ACL Rolling Review (2024, 2025), COLM (2025), ICLR (2025, 2026), NeurIPS (2025)

Workshop Reviewer: Reasoning and Planning for LLMs @ ICLR2025, LLM for Scientific Discovery @ COLM 2025

Secondary Reviewer: TACL (2024)

SKILLS

ML/LLM Programming

Pytorch, Huggingface, vLLM, OpenRLHF, veRL, asyncio, Ray

Languages

*Chinese (Native), English (Fluent), Japanese (Proficient; **JLPT N2**), Korean (Elementary)*