

Ziyu Yao

Dept of Computer Science & Engineering
The Ohio State University
2015 Neil Avenue
Columbus, OH 43210 USA

📧 ziyuyao.org ✉ yao.470@osu.edu 🐦 @ZiyuYao

Research Interests

My research interests lie in Natural Language Processing and Artificial Intelligence, particularly building natural language interfaces (such as semantic parsers and question answering systems) for knowledge discovery and problem solving. I am also passionate about applying NLP/AI techniques to domains such as Software Engineering and Healthcare.

Education

- 2017–present **Ph.D. Candidate**, The Ohio State University (OSU), Computer Science & Engineering. (Advisor: Prof. Huan Sun)
- 2015–2016 **M.Sc. Program**, The Ohio State University (OSU), Computer Science & Engineering.
- 2011–2015 **B.E.**, Beijing University of Posts and Telecommunications (BUPT), Communication Engineering.

Research Experience

- 2017–present **Research Assistant**, OSU (Advisor: Huan Sun).
- 2020 Summer **Research Intern**, LTI, Carnegie Mellon University (Mentor: Graham Neubig).
- 2018 Summer **Research Intern**, Microsoft Research, Redmond (Mentor: Yizhe Zhang, Bill Dolan).
- 2017 Summer **Research Intern**, Fujitsu Lab of America (Mentor: Mehdi Bahrami, Wei-Peng Chen).
- 2014–2015 **Research Intern**, Tsinghua University, Beijing (Mentor: Minlie Huang).

Awards and Honors

- 2021 **Graduate Student Research Award**, OSU, CSE.
- 2020 **Presidential Fellowship**, OSU.
- 2020 **Rising Star in EECS**, UC Berkeley.
- 2020 **Graduate Students Poster Award (Honorable Mention)**, CSE, OSU.
- 2019 **NSF Student Travel Award**, The Web Conference (WWW).
- 2019 **Graduate Students Poster Award (Honorable Mention)**, CSE, OSU.
- 2018 **Travel Grant for Graduate Cohort for Women, San Francisco**, CRA-W.
- 2015 **Beijing Outstanding Graduate Award**, Beijing, China.
- 2012, 2014 **First-class Scholarship (top 3%)**, BUPT, China.
- 2013 **Second-class Scholarship (top 5%)**, BUPT, China.

Publications

- **Ziyu Yao**, Frank Xu, Pengcheng Yin, Huan Sun, Graham Neubig, “Learning Structural Edits via Incremental Tree Transformations”, To appear at The Ninth International Conference on Learning Representations (**ICLR 2021**). [Paper, Code]
- **Ziyu Yao**, Yiqi Tang, Scott Wen-tau Yih, Huan Sun, Yu Su, “An Imitation Game for Learning Semantic Parsers from User Interaction”, 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2020, acceptance rate: 22.4%**). [Paper, Code]

- Xiang Yue, Xinliang Frederick Zhang, **Ziyu Yao**, Simon Lin, Huan Sun, “CliniQG4QA: Generating Diverse Questions for Domain Adaptation of Clinical Question Answering”, **Machine Learning for Health Workshop at NeurIPS 2020 (extended abstract track)**. [Paper, Code]
- **Ziyu Yao**, Yu Su, Huan Sun, Scott Wen-tau Yih, “Model-based Interactive Semantic Parsing: A Unified Formulation and A Text-to-SQL Case Study”, 2019 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2019, acceptance rate: 23.8%**). [Paper, Code]
- Boyuan Pan, Hao Li, **Ziyu Yao**, Deng Cai, Huan Sun, “Reinforced Dynamic Reasoning for Conversational Question Generation”, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (**ACL 2019, acceptance rate: 25.7%**). [Paper, Code]
- **Ziyu Yao**, Jayavardhan Reddy Peddamail, Huan Sun, “CoaCor: Code Annotation for Code Retrieval with Reinforcement Learning”, The Web Conference 2019 (**WWW 2019, acceptance rate: 18%**). [Paper, Code]
- **Ziyu Yao**, Xiujun Li, Jianfeng Gao, Brian Sadler, Huan Sun, “Interactive Semantic Parsing for If-Then Recipes via Hierarchical Reinforcement Learning”, The AAAI Conference on Artificial Intelligence 2019 (**AAAI 2019, acceptance rate: 16.2%; SPOTLIGHT**). [Paper, Code]
- **Ziyu Yao**, Yizhe Zhang, Xiujun Li, Jianfeng Gao, Michel Galley, Chris Brockett, Huan Sun, Bill Dolan, “IEC: Towards Interest-Eliciting Neural Conversational Agents”, **Manuscript 2019**, [Paper].
- Jayavardhan Reddy Peddamail, **Ziyu Yao**, Zhen Wang, Huan Sun, “A Comprehensive Study of StaQC for Deep Code Summarization”, **SIGKDD 2018 Deep Learning Day**. [Paper]
- **Ziyu Yao**, Daniel S. Weld, Wei-Peng Chen, Huan Sun, “StaQC: A Systematically Mined Question-Code Dataset from Stack Overflow”, The Web Conference 2018 (**WWW 2018, acceptance rate: 14.8%**). [Paper, Code]
- Li Zhao, Minlie Huang, **Ziyu Yao**, Rongwei Su, Yingying Jiang, Xiaoyan Zhu. “Semi-Supervised Multinomial Naive Bayes for Text Classification by Leveraging Word-Level Statistical Constraint”, The AAAI Conference on Artificial Intelligence 2016 (**AAAI 2016, acceptance rate: 25.8%**). [Paper]
- **Ziyu Yao**, Shouzhong Tu, Minlie Huang, Xiaoyan Zhu. “A Semi-Supervised Method for Filtering Chinese Spam Tweets”. **Journal of Chinese Information Processing 2016**, 30(5): 176-186. [Paper in Chinese]

Talks

- 2021 **Building Interactive Natural Language Interfaces**, Invited talk at Brandeis University (virtually). Host: Hongfu Liu.
- 2021 **Building Interactive Natural Language Interfaces**, OSU AI Club (virtually).
- 2020 **An Imitation Game for Learning Semantic Parsers from User Interaction**, EMNLP (virtually).
- 2020 **On Building an Interactive Natural Language Interface: An intelligent agent that can communicate, collaborate and learn**, Invited talk at Microsoft Research AI Breakthroughs Workshop (virtually). Invited by: Ahmed Hassan Awadallah, Alex Polozov.
- 2020 **Building Natural Language Interfaces to Programs**, Invited talk at VMware, Beijing (virtually). Host: Jin He, Kevin Song.

- 2020 **Towards Building Interactive and Collaborative Natural Language Interfaces**, Invited talk at NeuLab, Carnegie Mellon University (virtually). Host: Graham Neubig.
- 2020 **Learning Semantic Parsers from User Interaction**, CSE AI Seminar, OSU.
- 2019 **Towards Building Interactive and Collaborative Natural Language Interfaces**, Invited talk at Chair of Management Information System, ETH Zurich, Host: Stefan Feuerriegel.
- 2019 **CoaCor: Code Annotation for Code Retrieval with Reinforcement Learning**, WWW, San Francisco, USA.
- 2019 **Interactive Semantic Parsing for If-Then Recipes via Hierarchical Reinforcement Learning**, AAIL, Hawaii, USA.
- 2018 **Interactive Semantic Parsing for If-Then Recipes via Hierarchical Reinforcement Learning**, CSE AI Seminar, OSU.
- 2018 **StaQC: A Systematically Mined Question-Code Dataset from Stack Overflow**, WWW, Lyon, France.

Teaching and Mentoring Experience

Lectures and Panels.

- 2020 Fall **CSE Panel: Graduate School Experience and Expectation**, CSE, OSU. Panelist.
- 2020 Fall **Foundations of Speech and Language Processing**, CSE, OSU. Guest lecturer.
- 2018 Fall **Course Panel: Data Management in the Cloud**, CSE, OSU. Guest panelist.
- 2017 Spring **Introduction to Computer Programming in C++**, CSE, OSU. Student instructor.

(Co-)Mentored Students.

Jayavardhan Reddy Peddamail (Master Student at OSU; now at Google)

- Topic: Source code search and annotation.
- Publication: **WWW 2019, SIGKDD 2018 workshop**.

Varun Ranganathan Dhanvanth (Master Student at OSU; now at Facebook)

Heming Sun (Undergraduate Student at OSU)

Tianshu Zhang (PhD Student at OSU)

Leslie Zhou (Master Student at OSU)

Professional Activity

Organizing Committee.

NLP4Prog at ACL2021: The First Workshop on Natural Language Processing for Programming 2021

Program Committee or Reviewer (Conference).

<u>MSR</u> : Mining Software Repositories Conference	2021
<u>NAACL</u> : Annual Conference of the North American Chapter of the Association for Computational Linguistics	2021
<u>EMNLP</u> : Conference on Empirical Methods in Natural Language Processing	2018, 2020
<u>AACL</u> : Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics	2020
<u>AAAI</u> : The AAAI Conference on Artificial Intelligence	2020
<u>NLPCC</u> : Conference on Natural Language Processing and Chinese Computing	2020
<u>CoNLL</u> : Conference on Computational Natural Language Learning	2019

Secondary/External Reviewer:

<u>ACL</u> : Annual Meeting of the Association for Computational Linguistics	2018, 2020
<u>KDD</u> : SIGKDD Conference on Knowledge Discovery and Data Mining	2017, 2019, 2020
<u>NAACL</u> : Annual Conference of the North American Chapter of the Association for Computational Linguistics	2019
<u>ICDM</u> : International Conference on Data Mining	2016, 2018
<u>WWW</u> : International World Wide Web Conference	2017
<u>CIKM</u> : International Conference on Information and Knowledge Management	2016

Program Committee or Reviewer (Workshop).

<u>IntEx-SemPar at EMNLP 2020</u> : Interactive and Executable Semantic Parsing Workshop	2020
<u>NLI at ACL 2020</u> : Natural Language Interface Workshop	2020
<u>SRW at ACL 2019</u> : Student Research Workshop	2019

Membership.

Association for Computational Linguistics	2019-present
-------------------------------------------	--------------

Department Service.

Member of The Council of Graduate Women in Computer Science	2020-present
-------------------------------------------------------------	--------------

References

Huan Sun, Assistant Professor, Dept of Computer Science & Engineering, OSU.

- o Email: sun.397@osu.edu
- o Homepage: <http://web.cse.ohio-state.edu/~sun.397/>

Scott Wen-tau Yih, Research Scientist, Facebook AI Research.

- o Email: scottyih@fb.com
- o Homepage: <http://scottyih.org/>

Graham Neubig, Associate Professor, Language Technology Institute, School of Computer Science, Carnegie Mellon University.

- o Email: gneubig@cs.cmu.edu
- o Homepage: <http://www.phontron.com/>

Yu Su, Assistant Professor, Dept of Computer Science & Engineering, OSU.

- o Email: su.809@osu.edu
- o Homepage: <https://ysu1989.github.io/>