Ansong Ni

ansong.ni@yale.edu • http://niansong1996.github.io • +1 (412) 641-9684

RESEARCH INTEREST

I am generally interested in natural language processing and machine learning. My current research focus is **code generation with large language models (LLMs)**. More specifically, I am interested in these topics:

- Enhancing code LLMs with program execution (neuro-symbolic methods)
- Harmonizing natural language and code for planning, reasoning and interacting with humans
- Synthesizing large-scale, high-quality data for training next-gen code LLMs

EDUCATION

Yale University - Computer Science Department, New Haven, CT — Sep 2020 - Jul 2024 (Expected)
Ph.D. student in Computer Science | GPA: 4.0/4.0
- Advisor: Dragomir Radev, Arman Cohan

Carnegie Mellon University - School of Computer Science, Pittsburgh, PA — Aug 2018 - Dec 2019
Master of Science in Computer Science | GPA: 4.11/4.33
- Research Advisor: Graham Neubig

Nanjing University -Software Institute, Nanjing, China — Aug 2014 - Jun 2018
Bachelor of Engineering in Software Engineering | GPA: 4.41/5.0
- Research Advisor: Ming Li

INDUSTRIAL RESEARCH EXPERIENCE

Google DeepMind - Learning for Code Team, Mountain View, CA — Summer 2023
Research Intern | Hosts: Pengcheng Yin, Charles Sutton
- Improve the capability of LLMs to reason about program execution. [P1]

Meta AI - FAIR NLP Group, Menlo Park, CA — Summer 2022
Research Intern | Mentors: Victoria Lin, Sida Wang
- Execution-augmented program reranking for improving large language models on NL2Code tasks. [C14]

Microsoft Research - Deep Learning Group, Redmond, WA — Oct 2021 - Apr 2022
Part-time Research Collaborator | Managers: Chenglong Wang, Jeevana Priya Inala
- Learning from self-sample correct and partially-correct programs. [C13]

Microsoft Research - Deep Learning Group, Redmond, WA — Summer 2021
Research Intern | Managers: Alex Polozov, Chris Meek
- Improving code generation with program tracing. [C13]

Allen Institute for AI - AllenNLP Team, Seattle, WA — Summer 2020
Research Intern | Managers: Pradeep Dasigi, Matt Gardner
- Mitigating false-negative context in multi-document question answering. [C7]

Microsoft Research Asia - Software Analytics Group, Beijing, China — Jun 2017 – Dec 2017
Research Intern | Manager: Shi Han
- Automatic mining of insights from multi-dimensional data. [link]

ACADEMIC RESEARCH EXPERIENCE

Yale University - LILY Lab, New Haven, CT — Sep 2020 - Current
Research Assistant | Advisor: Dragomir Radev
- Semantic parsing and code generation. [C10, C13, C14, J1, W1, P1]
- Long document and dialogue summarization. [C5, C6, C8, C9, C11]

Carnegie Mellon University - Institute for Software Research, Pittsburgh, PA — Spring 2020
Research Assistant | Advisor: Claire Le Goues, Ruben Martins
- Combining NLP and program synthesis for automatic API migration between different libraries. [C4]

Carnegie Mellon University - Language Technology Institute, Pittsburgh, PA — Spring 2019
Research Associate | Advisor: Graham Neubig

- Active learning for weakly-supervised semantic parsing. [C3]

Nanjing University - LAMDA Group, Nanjing, China — 2017-2018
Research Assistant | Advisor: Ming Li

- Learning to predict failures in continuous integration. [C1, C2]

PUBLICATIONS

[Preprints] *(*: denotes equal contribution)*

[P1] **NExT: Teaching Large Language Models to Reason about Code Execution**
Ansong Ni, Miltos Allamanis, Arman Cohan, Yinlin Deng, Kensen Shi, Charles Sutton, Pengcheng Yin
In Submission to ICML'24 Long Paper (arxiv available soon)

[Journal Papers]

[J1] **L2CEval: Evaluating Language-to-Code Generation Capabilities of Large Language Models**
Ansong Ni, Pengcheng Yin, Yilun Zhao, Martin Riddell, Troy Feng, Rui Shen, Stephen Yin, Ye Liu, Semih Yavuz, Caiming Xiong, Shafiq Joty, Yingbo Zhou, Dragomir Radev, Arman Cohan
(To appear) Transactions of the Association for Computational Linguistics (TACL'24) Long Paper [pdf]

[Conference Papers]

[C14] **LEVER: Learning to Verify Language-to-Code Generation with Execution**
Ansong Ni, Srini Iyer, Dragomir Radev, Ves Stoyanov, Wen-tau Yih, Sida I. Wang, Victoria X. Lin**
The 40th International Conference on Machine Learning (ICML'23) Long Paper [pdf]

[C13] **Learning Math Reasoning from Self-Sampled Correct and Partially-Correct Solutions**
Ansong Ni, Jeevana Priya Inala, Chenglong Wang, Oleksandr Polozov, Christopher Meek, Dragomir Radev, Jianfeng Gao
The 11th International Conference on Learning Representations (ICLR'23) Long Paper [pdf]

[C11] **Leveraging Locality in Abstractive Text Summarization**
Yixin Liu, Ansong Ni, Linyong Nan, Budhaditya Deb, Chenguang Zhu, Ahmed H. Awadallah, Dragomir Radev
The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP'22) Long Paper [pdf]

[C10] **UnifiedSKG: Unifying and Multi-Tasking Structured Knowledge Grounding with Text-to-Text Language Models**
Tianbao Xie, Chen Henry Wu*, Peng Shi, Ruiqi Zhong, Torsten Scholak, Michihiro Yasunaga, Chien-Sheng Wu, Ming Zhong, Pengcheng Yin, Sida I. Wang, Victor Zhong, Bailin Wang, Chengzu Li, Connor Boyle, Ansong Ni, Ziyu Yao, Dragomir Radev, Caiming Xiong, Lingpeng Kong, Rui Zhang, Noah A. Smith, Luke Zettlemoyer, and Tao Yu*
The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP'22) Long Paper [pdf]

[C9] **DYLE: Dynamic Latent Extraction for Abstractive Long-Input Summarization**
Ziming Mao, Chen Henry Wu, Ansong Ni, Yusen Zhang, Rui Zhang, Tao Yu, Budhaditya Deb, Chenguang Zhu, Ahmed Hassan Awadallah, and Dragomir Radev
The 60th Annual Meeting of the Association for Computational Linguistics (ACL'22) Long Paper [pdf]

[C8] **Summ^N: A Multi-Stage Summarization Framework for Long Input Dialogues and Documents**
Yusen Zhang, Ansong Ni, Ziming Mao, Chen Henry Wu, Chenguang Zhu, Budhaditya Deb, Ahmed Hassan Awadallah, Dragomir Radev, and Rui Zhang
The 60th Annual Meeting of the Association for Computational Linguistics (ACL'22) Long Paper [pdf]

[C7] **Mitigating False-Negative Contexts in Multi-document Question Answering with Retrieval Marginalization**
Ansong Ni, Matt Gardner, and Pradeep Dasigi
The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP'21) Long Paper [pdf]

[C6] **An Exploratory Study on Long Dialogue Summarization: What Works and What's Next**
Yusen Zhang, Ansong Ni*, Tao Yu, Rui Zhang, Chenguang Zhu, Budhaditya Deb, Asli Celikyilmaz, Ahmed Hassan Awadallah, and Dragomir Radev*
The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP'21) Findings [pdf]

[C5] **SummerTime: Text Summarization Toolkit for Non-experts**
Ansong Ni, Zhangir Azerbayev, Mutethia Mutuma, Troy Feng, Yusen Zhang, Tao Yu, Ahmed Hassan Awadallah, and Dragomir Radev
The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP'21) Demo Track [pdf]

[C4] **SOAR: A Synthesis Approach for Data Science API Refactoring**
Ansong Ni, Daniel Ramos*, Aidan Yang, Ines Lynce, Vasco Manquinho, Ruben Martins, and Claire Le Goues*
The 43th International Conference on Software Engineering (ICSE'21) Long Paper [pdf]

[C3] **Merging Weak and Active Supervision for Semantic Parsing**
Ansong Ni, Pengcheng Yin, and Graham Neubig
The 34th AAAI Conference on Artificial Intelligence (AAAI'20) Long Paper [pdf]

[C2] **ACONA: Active Online Model Adaptation for Predicting Continuous Integration Build Failures**
Ansong Ni and Ming Li
The 40th International Conference on Software Engineering (ICSE'18) Poster Track

[C1] **Cost-effective Build Outcome Prediction Using Cascaded Classifier**
Ansong Ni and Ming Li
The 14th International Conference on Mining Software Repositories (MSR'17) Mining Challenge

[Workshop Papers]

[W1] **Explicit Knowledge Transfer for Weakly-Supervised Code Generation**
Zhangir Azerbayev, Ansong Ni, Hailey Schoelkopf, Dragomir Radev
Deep Learning For Code (DL4C) Workshop @ ICLR'23 [pdf]

TALKS & PRESENTATIONS

Foundation Models for Code and Math
- Guest Lecture @ Yale CPSC 488/588 "AI Foundation Models" — Dec 2023

Enhancing Language Models for Program Synthesis using Execution
- Invited Talk @ MIT CSAIL — Mar 2023
- Invited Talk @ HKUST CSE — Mar 2023
- Invited Talk @ UT Austin - TAUR Lab — Apr 2023

LEVER: Learning to Verify Language-to-Code Generation with Execution
- Paper Presentation @ Google Brain Program Synthesis Reading Group — Jul 2023

Learning from Self-Sampled Correct and Partially-Correct Programs
- Paper Presentation @ Meta AI Reading Group — Jun 2022

Merging Weak and Active Supervision for Semantic Parsing
- Oral Presentation @ AAAI Conference — Feb 2020

PROFESSIONAL SERVICES

- Program Committee / Reviewer
  - COLM 2024
  - ICLR 2024
  - ICML 2023, 2024
  - NeurIPS 2022, 2023
  - ACL 2023
  - EMNLP 2022
  - ACL ARR 2021-2023
  - DL4C Workshop @ ICLR 2023
  - SUKI Workshop @ NAACL 2022
  - IntEx-SemPar Workshop @ EMNLP 2020

HONORS & AWARDS

- AWS Cloud Credits for Research Program ($10,000) — Yale University, Apr 2023
- University Nominee for Google Fellowship — Yale University, Oct 2022
- University Nominee for Microsoft Fellowship — Yale University, July 2022
- CS Department Fellowship — Yale University, Aug 2020
- Outstanding Graduate — Nanjing University, Jun 2018

TEACHING & MENTORING

- Teaching Fellow:
  - CPSC 477/577 Natural Language Processing
  - CPSC 482/582 Topics in Applied Machine Learning
- Mentored undergraduates (projects) at Yale:
  - Ziming Mao (long-input summarization [C9]) → CS PhD @ UC Berkeley
  - Zhangir Azerbayev (code generation [W1]) → CS PhD @ Princeton
  - Hailey Schoelkopf (code generation [W1]) → Research Scientist @ Eleuther AI

PROGRAMMING SKILLS

- Programming Languages (ranked by proficiency): Python, Shell, CUDA, C/C++, Java
- Frameworks and Tools: PyTorch, Vim
- Featured Repositories:
  - SummerTime (Text summarization toolkit for non-experts, 250+ stars) [Link]
  - LEVER (Experiment code for the LEVER paper) [Link]