

Lunyu Nie

Email: nlx20@mails.tsinghua.edu.cn | Homepage: flitternie.github.io

EDUCATION

- | | |
|--|----------------|
| Tsinghua University , Beijing, China | 2020 – Present |
| M.S. in Computer Science and Technology | |
| - CGPA: 3.76 / 4.00 | |
| - Advisor: Professor Jidong Zhai and Professor Juanzi Li | |
| The Chinese University of Hong Kong , Hong Kong | 2015 – 2020 |
| B.B.A. with Honours in Accountancy, Second Major in Computer Science | |
| - CGPA: 3.60 / 4.00, Computer Science GPA: 3.70 / 4.00 | |
| - Advisor: Professor Helen M. Meng and Professor Wai Lam | |

PUBLICATIONS

- [1] GraphQ IR: Unifying the Semantic Parsing of Graph Query Languages with One Intermediate Representation.
Lunyu Nie, Shulin Cao, Jiaxin Shi, Qi Tian, Lei Hou, Juanzi Li, Jidong Zhai.
To appear in *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022.
- [2] Guiding the PLMs with Semantic Anchors as Intermediate Supervision: Towards Interpretable Semantic Parsing.
Lunyu Nie, Jiuding Sun, Yanling Wang, Lun Du, Han Shi, Dongmei Zhang, Lei Hou, Juanzi Li, Jidong Zhai.
Submitted to the *AAAI Conference on Artificial Intelligence (AAAI)*, 2022.
- [3] KQA Pro: A Dataset with Explicit Compositional Programs for Complex Question Answering over Knowledge Base.
Shulin Cao, Jiaxin Shi, Liangming Pan, **Lunyu Nie**, Yutong Xiang, Lei Hou, Juanzi Li, Hanwang Zhang, Bin He.
In *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2022.
- [4] CoreGen: Contextualized Code Representation Learning for Commit Message Generation.
Lunyu Nie, Cuiyun Gao, Zhicong Zhong, Wai Lam, Yang Liu, Zenglin Xu.
In *Neurocomputing*, 2021.
- [5] Code Structure Guided Transformer for Source Code Summarization.
Shuzheng Gao, Cuiyun Gao, Yulan He, Jichuan Zeng, **Lunyu Nie**, Xin Xia, Michael Lyu.
In *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 2022.
- [6] Unstructured Knowledge Access in Task-oriented Dialog Modeling using Language Inference, Knowledge Retrieval and Knowledge-Integrative Response Generation.
Mudit Chaudhary, Borislav Dzodzo, Sida Huang, Chun Hei Lo, Mingzhi Lyu, **Lunyu Nie**, Jinbo Xing, Tianhua Zhang, Xiaoying Zhang, Jingyan Zhou, Hong Cheng, Wai Lam, Helen Meng. (Authors ordered alphabetically)
In *AAAI-21 Workshop on 9th Dialog System Technology Challenge (DSTC9)*, 2021.
- [7] ATOM: Commit Message Generation Based on Abstract Syntax Tree and Hybrid Reranking.
Shangqing Liu, Cuiyun Gao, Sen Chen, **Lunyu Nie**, Yang Liu.
In *IEEE Transactions on Software Engineering (TSE)*, 2020.

RESEARCH AND INDUSTRY EXPERIENCE

- | | |
|--|-------------|
| Microsoft Research Asia, Software Analytics Group | Summer 2022 |
| <i>Research Intern</i> , with Yanlin Wang and Lun Du | |
| The Chinese University of Hong Kong, Human-Computer Communications Laboratory | 2020 |
| <i>Undergraduate Research Assistant</i> , with Professor Helen M. Meng | |
| The Chinese University of Hong Kong, Text-Mining Group | 2019 – 2020 |
| <i>Undergraduate Research Assistant</i> , with Professor Wai Lam | |
| Chinese Academy of Sciences, Institute of Computing Technology | Summer 2018 |
| <i>Research Intern</i> , with Professor Yungang Bao | |

RESEARCH INTERESTS

Fields	Semantic Parsing, Code Representation Learning, Knowledge Reasoning, Transfer Learning
Methods	Deep Learning, Programming Language Theory, Reinforcement Learning

SELECTED RESEARCH PROJECTS

Unified Intermediate Representation for Graph Query Languages	2021 - Present
<ul style="list-style-type: none">- Proposed <i>GraphQ IR</i>, an intermediate representation that aims to bridge the semantic gap between natural language and graph query languages.- Developed <i>GraphQ Trans</i>, a source-to-source compiler that for the first time in the field supports translation among multiple graph query languages (SPARQL, Cypher, Gremlin, Lambda-DCS, KoPL, etc.).- Achieved state-of-the-art performance on several semantic parsing benchmarks <i>Overnight</i>, <i>KQA Pro</i>, <i>GrailQA</i> and <i>MetaQA-Cypher</i> with promising generalizability (11% accuracy improvement) under non-IID (compositional generalization and zero-shot) and low-resource settings.	

Automatic Code Performance Optimization with Deep Learning	2020 - 2021
<ul style="list-style-type: none">- Built up a Python performance bug dataset by collecting the bug-fixing commits from open-source communities & developed a system for automatic Python performance downgrading detection with 94.3% recall rate.- Developed a code pretrained model that predicts compiler optimization heuristics and achieved state-of-the-art performance with at most $3.54\times$ performance speedup on heterogeneous compute device mapping task.	

Code Representation Learning for Commit Message Generation & Code Summarization	2018 - 2020
<ul style="list-style-type: none">- Proposed <i>CoreGen</i>, a novel code representation learning method that exploits the contextual information behind the code changes via self-supervised learning with at least 28.18% BLEU-4 score improvement.- Developed <i>ATOM</i>, a commit message generation framework that explicitly incorporates the abstract syntax trees for representing the code changes and ensembles retrieval & generative methods via hybrid reranking.- Developed <i>SG-Trans</i>, a code summarization model that utilizes hierarchical code structures, i.e., local symbolic information and global syntactic structure (e.g., data flow) as the inductive biases in self-attention.	

Task-oriented Dialog Modelling with Unstructured Knowledge Access	2020
<ul style="list-style-type: none">- Developed a task-oriented dialogue system that incorporates external unstructured knowledge for answering out-of-API-coverage user queries through knowledge retrieval and ensembled response generation.- Outperformed the baseline with a 58.77% BLEU-4 score improvement, ranked as the global finalist in <i>the Ninth Dialog System Technology Challenge</i> (DSTC9) Track 1 with paper presented in <i>AAAI-21</i> workshop.	

SCHOLARSHIPS & AWARDS

- 2022, First-class Scholarship for Hong Kong & Macau Students, Ministry of Education (Highest scholarship awarded by the Chinese government for Hong Kong Students, <0.1%)
- 2022, Outstanding Social Work Award, Tsinghua University
- 2021, Weihai Talent Scholarship, Tsinghua University (Top 10% in Department)
- 2018 ~ 2020, Faculty Dean's List, Chinese University of Hong Kong (Top 10% in Faculty)
- 2018 ~ 2019, Wu Yee Sun College Master's List, Chinese University of Hong Kong (1st per major/class)
- 2017, EY Scholarship (2 awardees/year)

EXTRACURRICULAR ACTIVITIES

Student Counsellor, Graduate Affairs Office of Tsinghua University	2021 – 2022
Provided guidance and organized activities for Tsinghua's Hong Kong, Macau, and Taiwan postgraduates.	

Debater, The Chinese University of Hong Kong Debate Team	2015 - 2018
Represented CUHK in international debate competitions and won prizes, e.g., Area Champion in Chinese Debate World Cup 2019, Champion in Hong Kong Inter-collegiate Mandarin Debate Competition 2018, etc.	

LANGUAGE SKILLS

TOEFL iBT	110/120 (Reading 30, Listening 29, Speaking 24, Writing 27)
GRE	329/340 + 4.0/6.0 (Verbal 159, Quantitative 170, Analytical Writing 4.0)