Qian Liu(刘乾)

liuqian@sea.com · https://siviltaram.github.io · SivilTaram

Education

Beihang University 2017.9 – 2022.6

Ph.D. in Computer Science and Engineering ♣ Joint Program with Microsoft Research Asia
Ph.D. Supervisor: Qinping Zhao & Bin Zhou ♣ Co-advisor: Jian-Guang Lou

Beihang University 2013.9 – 2017.6

B.S. in Computer Science and Technology, Ranking 7 / 233

Experience

Sea AI Lab, Singapore 2022.7 – Present

Research Scientist at *Natural Language Processing* Group ♢ Leader: Shuicheng Yan

Microsoft Research Asia, China 2018.2 – 2022.6

Research Intern at *Data, Knowledge and Intelligence* Group ♢ Leader: Jian-Guang Lou & Bei Chen

Microsoft Research Asia, China 2016.7 – 2017.8

Research Intern at *Big Data Mining* Group ♢ Leader: Zaiqing Nie & Yohn Cao

Microsoft Asia-Pacific R&D Group, China 2016.5 – 2016.7

Development Intern at *Operating System* Group ♢ Leader: Xin Zou & Jianwei Zhu

Research Interests

- **Semantic Parsing**: *Focusing on building generalizable, interactive and interpretable text-to-SQL systems.*
- **Reinforcement Learning**: *Focusing on applying reinforcement learning on various NLP applications.*

Selected Projects

Text-to-Formula Overnight 2021.1 – 2021.6

This project is to build a text-to-formula semantic parser for Microsoft PowerApp. To make the parser work with low or even no parallel data, we employ the technique of data synthesis to leash the power of pre-trained language models (e.g., GPT-3) on semantic parsing. The project has been announced in Microsoft Build.

Compositional Generalization 2020.2 – Present

Compositional generalization is a basic but essential intellective capability of human beings. However, existing neural models have been proven to be extremely deficient in such a capability. To tackle it, we propose a novel memory-augmented neural model. Specifically, our model is the first neural model to pass all compositional challenges addressed by previous works without extra resources. As one of the main contributor, I contribute the original idea, the model structure and the training protocol.

Conversational Text-to-SQL 2018.2 – Present

This ongoing research project is born out of demand from Microsoft products, which aims to develop a conversational text-to-SQL system to help common users to analyze tabular data through natural language conversations. To improve the accuracy of our system, we explore various context modeling techniques (query rewriting, turn-level encoding, tree-based copy mechanism). As the owner of this project, I am responsible for algorithm design, system implementation, and working on experiments.

Selected Publications

- Xinyu Pi*, **Qian Liu***, Bei Chen, Morteza Ziyadi, Zeqi Lin, Yan Gao, Qiang Fu, Jian-Guang Lou, Weizhu Chen, Reasoning Like Program Executors. Under Review (* = equal contribution)
- **Qian Liu**, Bei Chen, Jiaqi Guo, Morteza Ziyadi, Zeqi Lin, Weizhu Chen, Jian-Guang Lou, TAPEX: Table Pre-training via Learning a Neural SQL Executor. In *International Conference on Learning Representations* 2022 (ICLR-2022, **Highest rating in the initial review**)
- Jiaqi Guo, Ziliang Si, Yu Wang, **Qian Liu**, Ming Fan, Jian-Guang Lou, Zijiang Yang, Ting Liu, CHASE: A Large-Scale and Pragmatic Chinese Dataset for Cross-Database Context-Dependent Text-to-SQL. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics* (ACL-2021, **Oral**)
- **Qian Liu***, Dejian Yang*, Jiahui Zhang*, Jiaqi Guo, Bin Zhou, Jian-Guang Lou, Awakening Latent Grounding from Pretrained Language Models for Semantic Parsing. In *Findings of the Association for Computational Linguistics: ACL 2021* (ACL-2021 Findings, * = equal contribution)
- Shuang Chen*, **Qian Liu***, Zhiwei Yu*, Chin-Yew Lin, Jian-Guang Lou, Feng Jiang, ReTraCk: A Flexible and Efficient Framework for Knowledge Base Question Answering. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (ACL-2021 Demo, * = equal contribution)
- Yuntao Li, Bei Chen, **Qian Liu**, Yan Gao, Jiang-Guang Lou, Yan Zhang, Dongmei Zhang, Keep the Structure: A Latent Shift-Reduce Parser for Semantic Parsing. In *30th International Joint Conference on Artificial Intelligence* (IJCAI-2021)
- **Qian Liu***, Shengnan An*, Jian-Guang Lou, Bei Chen, Zeqi Lin, Yan Gao, Bin Zhou, Nanning Zheng, Dongmei Zhang, Compositional Generalization by Learning Analytical Expressions. In *Advances in Neural Information Processing Systems 34* (NeurIPS-2020, **Spotlight**, * = equal contribution)
- **Qian Liu**, Bei Chen, Jian-Guang Lou, Bin Zhou and Dongmei Zhang, Incomplete Utterance Rewriting as Semantic Segmentation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing* (EMNLP-2020)
- Jiaqi Guo, **Qian Liu**, Jian-Guang Lou, Zhenwen Li, Xueqing Liu, Tao Xie and Ting Liu, Benchmarking Meaning Representations in Neural Semantic Parsing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing* (EMNLP-2020)
- Yuntao Li, Bei Chen, **Qian Liu**, Yan Gao, Jian-Guang Lou, Yan Zhang and Dongmei Zhang, "What Do You Mean by That?" A Parser-Independent Interactive Approach for Enhancing Text-to-SQL. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing* (EMNLP-2020)
- **Qian Liu**, Bei Chen, Jiaqi Guo, Jian-Guang Lou, Bin Zhou and Dongmei Zhang, How Far are We from Effective Context Modeling? An Exploratory Study on Semantic Parsing in Context. In *29th International Joint Conference on Artificial Intelligence* (IJCAI-2020)
- **Qian Liu**, Yihong Chen, Bei Chen, Jian-Guang Lou, Zixuan Chen, Bin Zhou, Dongmei Zhang, You Impress Me: Dialogue Generation via Mutual Persona Perception. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (ACL-2020)
- **Qian Liu**, Bei Chen, Haoyan Liu, Lei Fang, Jian-Guang Lou, Bin Zhou, Dongmei Zhang, A Split-and-Recombine Approach for Follow-up Query Analysis. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing* (EMNLP-2019)
- **Qian Liu**, Bei Chen, Jian-Guang Lou, Ge Jin, Dongmei Zhang, FANDA: A Novel Approach to Perform Follow-up Query Analysis. In *33th AAAI Conference on Artificial Intelligence* (AAAI-2019)

Skills

- Programming Languages: Python > C# > JavaScript > HTML/CSS = C++
- Deep Learning Tools: FairSeq > AllenNLP = PyTorch > Transformers > Tensorflow

Honors & Awards

Baidu Scholarship Top 20 Worldwide	2020
Graduate Student National Scholarship	2019, 2021
Beijing Outstanding Graduate Awards	2017