

QIAN LIU(刘乾)

✉ liuqian@sea.com · ☎ (+86) 188-1116-0363 · 🔗 <https://siviltaram.github.io> · 🌐 SivilTaram

👤 EXPERIENCE

Sea AI Lab, Singapore 2022.7 – Present

Research Scientist & Team Lead, *Natural Language Processing* Group

Microsoft Research Asia, China 2018.2 – 2022.6

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

Microsoft Research Asia, China 2016.7 – 2017.8

Research Intern, *Big Data Mining* Group ◇ Leader: Zaiqing Nie & Yohn Cao

🎓 EDUCATION

Beihang University 2017.9 – 2022.6

Ph.D. in Computer Science and Engineering ♣ Joint Program with Microsoft Research Asia

Ph.D. Supervisor: Qinqing Zhao

Beihang University 2013.9 – 2017.6

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

💡 RESEARCH INTERESTS

- **Semantic Parsing:** *Building generalizable, interactive and interpretable text-to-SQL systems.*
- **Language Reasoning:** *Improving fundamental reasoning capabilities of language models.*

📌 SELECTED PROJECTS

Synthetic Pre-training for Language Reasoning 2021.7 – Present

Language model pre-training has achieved great success via leveraging large-scale textual data. However, it is still very challenging for language models to perform reasoning because of the difficulty of mining large-scale reasoning-oriented natural language corpus. In this ongoing research, we have developed a series of methods to demonstrate that synthetic programs can also be a good source for reasoning pre-training. Experimental results show that our method can achieve significant improvements on tabular reasoning, numerical reasoning and spatial reasoning across various benchmarks. As the leader of this project, I am responsible for algorithm design, system implementation, and working on experiments.

📖 SELECTED PUBLICATIONS

- Chengsong Huang, Qian Liu, Bill Yuchen Lin, Tianyu Pang, Chao Du, Min Lin, [LoraHub: Efficient Cross-Task Generalization via Dynamic LoRA Composition](#). (Preprint, 2023)
- 🌸 BigCode, [StarCoder: may the source be with you!](#) (Preprint, 2023, **2.3K** Like in [HuggingFace Hub](#))
- Jianan Zhao, Meng Qu, Chaozhuo Li, Hao Yan, **Qian Liu**, Rui Li, Xing Xie, Jian Tang, [Learning on Large-scale Text-attributed Graphs via Variational Inference](#). In *International Conference on Learning Representations 2023 (ICLR-2023 Oral)*
- Bill Yuchen Lin, Chengsong Huang, **Qian Liu**, Wenda Gu, Sam Sommerer, Xiang Ren, [On Grounded Planning for Embodied Tasks with Language Models](#). In *37th AAAI Conference on Artificial Intelligence (AAAI-2023 Oral)*
- Qi Shi, **Qian Liu**, Bei Chen, Yu Zhang, Ting Liu, Jian-Guang Lou, [LEMON: Language-Based Environment Manipulation via Execution-Guided Pre-training](#). In *Findings of the Association for Computational Linguistics: EMNLP 2022 (EMNLP-2022 Findings)*

- Xinyu Pi*, **Qian Liu***, Bei Chen, Morteza Ziyadi, Zeqi Lin, Yan Gao, Qiang Fu, Jian-Guang Lou, Weizhu Chen, [Reasoning Like Program Executors](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP-2022 Oral, MLADS-2022 Distinguished Contribution Award, * = 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 1st Round)*
- 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)*
- **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

Beihang Outstanding Doctoral Student Awards	2022
Ph.D. Academic Excellence Foundation	2021
Baidu Scholarship Top 20 Worldwide	2020
Graduate Student National Scholarship	2019, 2021
Beijing Outstanding Graduate Awards	2017