Sitao Cheng

sitao.cheng@uwaterloo.ca | https://sitaocheng.github.io/

Research Interest

I am passionate about Language Agents, Reasoning and Retrieval-augmented Generation. I have worked on reinforcement learning (*e.g.*, mechanistic analysis and automatic reward modeling) and knowledge-intensive reasoning (*e.g.*, structured and unstructured data). Currently, I focus on compositional generalization in real-world tasks.

Education

- **University of Waterloo** — *09.2025 - 01.2029 (anticipated)*
  Ph.D. in Computer Science - advised by Prof. Victor Zhong — Waterloo, Canada
- **Nanjing University** — *09.2021 - 06.2024*
  M.S. in Computer Science and Technology - advised by Prof. Yuzhong Qu — Nanjing, China
- **University of Electronic Science and Technology of China** — *09.2017 - 06.2021*
  B.E. in Software Engineering - GPA: 3.98 / 4.00 (Top 3) — Chengdu, China

Publications

*Equal Contribution

Preprints.

[1] **From Atomic to Composite: Reinforcement Learning Enables Generalization in Complementary Reasoning**. [Paper] [Code]
Sitao Cheng, Xunjian Yin, Ruiwen Zhou, Yuxuan Li, Xinyi Wang, Liangming Pan, William Yang Wang, Victor Zhong

[2] **Differentiable Evolutionary Reinforcement Learning**. [Paper] [Code]
Sitao Cheng*, Tianle Li*, Xuhan Huang*, Xunjian Yin, Difan Zou

Selected Conference paper.

[1] **Understanding the Interplay between Parametric and Contextual Knowledge for Large Language Models**. ACL KnowFM Workshop (Oral), 2025. [Paper] [Code]
Sitao Cheng, Liangming Pan, Xunjian Yin, Xinyi Wang, William Yang Wang

[2] **Call Me When Necessary: LLMs can Efficiently and Faithfully Reason over Structured Environments**. ACL findings, 2024. [Paper] [Code]
Sitao Cheng, Ziyuan Zhuang, Yong Xu, Fangkai Yang, Chaoyun Zhang, Xiaoting Qin, Xiang Huang, Ling Chen, Qingwei Lin, Dongmei Zhang, Saravan Rajmohan, Qi Zhang

[3] **Reverse Language Model**. ACL, 2026 [Paper]
Xunjian Yin, **Sitao Cheng**, Yuxi Xie, Xinyu Hu, Li Lin, Xinyi Wang, Liangming Pan, William Yang Wang, Xiaojun Wan

[4] **QueryAgent: a Reliable and Efficient Reasoning Framework with Environmental Feedback-based Self-Correction**. ACL (Oral), 2024. [Paper] [Code]
Xiang Huang*, **Sitao Cheng***, Shanshan Huang, Jiayu Shen, Yong Xu, Chaoyun Zhang, Yuzhong Qu

[5] **MarkQA: a Large Scale KBQA Dataset with Numerical Reasoning**. EMNLP, 2023. [Paper] [Code]
Xiang Huang, **Sitao Cheng**, Yuheng Bao, Shanshan Huang, Yuzhong Qu

[6] **Question Decomposition Tree for Answering Complex Questions over Knowledge Bases**. AAAI (Oral), 2023. [Paper] [Code]
Xiang Huang, **Sitao Cheng**, Yiheng Shu, Yuheng Bao, Yuzhong Qu

[7] **EfficientRAG: Efficient Retriever for Multi-Hop Question Answering**. EMNLP, 2024. [Paper] [Code]
Ziyuan Zhuang*, Zhiyang Zhang*, **Sitao Cheng**, Fangkai Yang, Jia Liu, Shujian Huang, Qingwei Lin, Saravan Rajmohan, Dongmei Zhang, Qi Zhang

[8] **Disentangling Memory and Reasoning Ability in Large Language Models**. ACL, 2025. [Paper] [Code]
Mingyu Jin, Weidi Luo, **Sitao Cheng**, Xinyi Wang, Wenyue Hua, Ruixiang Tang, William Yang Wang, Yongfeng Zhang

[9] **RuleArena: A Benchmark for Rule-Guided Reasoning with LLMs in Real-World Scenarios**. ACL, 2025. [Paper] [Paper]
Ruiwen Zhou, Wenyue Hua, Liangming Pan, **Sitao Cheng**, Xiaobao Wu, En Yu, William Yang Wang

Research Experience

- **University of Waterloo (R2L Lab)** *09.2025 - Now*
  ***Advisor**: Prof. Victor Zhong. **Role**: Ph.D. Student* Waterloo, Canada
  - **Topic 1**: Evaluation and understanding of RL in compositional generalization. Exploration of RL as a synthesizer or amplifier with controlled experiments. Investigation of the condition of RL-driven generalization.
    - Findings: RL is genuinely a synthesizer instead of a probability amplifier, under a strict condition that the base model captures sufficient atomic skills.
    - Results: One Preprint.
  - **Topic 2**: A differentiable evolutionary training framework for reinforcement learning (DERL).
    - Description: Introduce a bi-level evolutionary training framework with a meta-optimizer to automatically generate reward functions for RL learning from validation signals. Introduce a bi-level evolutionary loop to jointly train the meta-optimizer and policy model.
    - Results: One Preprint.
- **University of California, Santa Barbara (NLP Group)** *07.2024 - 06.2025*
  ***Advisor**: Prof. William Wang. **Role**: Visiting Research Scholar* Santa Barbara, U.S.A
  - **Topic**: Understanding how effective LLMs leverage parametric knowledge when contextual knowledge is given.
    - Findings: LLMs consistently suppress their own knowledge given the context, regardless of models, knowledge types, the relations between two knowledge sources, and various levels of instructions.
    - Results: One publication on ACL 2025. Four other collaboration projects published.
- **Microsoft Research Asia (DKI Group)** *10.2023 - 06.2024*
  ***Advisor**: Yong Xu, Fangkai Yang, Chaoyun Zhang. **Role**: Research Intern & Mentor of Junior Interns* Beijing, China
  - **Topic 1**: LLMs reasoning framework over structured environments with retrieval-augmented generation (Readi) or neural symbolic reasoning (QueryAgent).
    - Description: Explore how LLMs can reason both efficiently and faithfully on large-scaled and heterogeneous structured environments? Adopt LLMs to either directly maintain a reasoning path (Readi), or step-by-step build a query (QueryAgent), both incorporating pertinent information for correction.
    - Results: Two publications on ACL 2024.
  - **Topic 2**: Efficient iterative retrieval with soley encoder-based models (EfficientRAG) and a new data organization paradigm (THREAD) for RAG systems.
    - Description: Model the link between the chunks. Leverage strong understanding ability of LLMs to reason the link between chunks. Design novel retrieval methods for smaller encoder-based models (EfficientRAG) and re-organize the documents (Thread), to model such link.
    - Results: One publication on EMNLP 2024. One publication on EMNLP 2025.
  - **Topic 3**: LLM-based Personalized AI Assistant with Structured Knowledge Graphs.
    - Results: One submission on CHI.
- **Nanjing University (Websoft Lab)** *09.2021 - 06.2024*
  ***Advisor**: Prof. Yuzhong Qu. **Role**: Student Researcher* Nanjing, China
  - **Topic 1**: Step-by-step query building (**QueryAgent**) with self-correction based on environmental feedback.
    - Description: Introduce a functional tool-set with environmental feedback and a zero-shot correction method for both reliability and efficiency.
    - Results: One publication on ACL 2024.
  - **Topic 2**: A KBQA benchmark (**MarkQA**) requiring both multi-hop and numerical reasoning ability.
    - Description: Propose NR-KBQA to challenge both reasoning ability over knowledge bases. Automatically build a large-scale dataset (MarkQA). Design PyQL query, a function toolset able to seamless SPARQL conversion, as symbolic reasoning steps, alleviating labeling burden.
    - Results: One publication on EMNLP 2023.
  - **Topic 3**: A question decomposition method (**QDT**) for better multi-hop reasoning over knowledge bases.
    - Description: Propose a serializable tree-based structure (QDT) to represent complex questions, which can sufficiently split questions with complex structures.
    - Results: One publication on AAAI 2023.

Honors and Awards

- **Layer 6 – TD Graduate Scholarship in Data & AI For Fall 2025**–UWaterloo
- **MCM/ICM H Prize, Outstanding Student of Sichuan Province, Outstanding Student Award**–NJU, UESTC

Others

- **Academic Services:** ACL volunteers, ARR Rolling Reviewers
- **Professional Skills:** Popular NLP models (LLM applications, Transformers, attention mechanism, etc.), Pytorch, C++, LaTex, Python, SQL
- **Interests:** Body building (over 6x body weight in the Big 3) , Basketball (member of department team), Swim
- **Social Service:** I serve as a personal assistant for a senior impressionist artist in UC Santa Barbara.