Sitao Cheng

sitao.cheng@uwaterloo.ca | sitaocheng.github.io

Research Interest

I am passionate about continual learning of language agents for efficient generalization. I have worked on long-horizon reasoning (*e.g.,* efficient and structured learning of experience), reinforcement learning (*e.g.,* conditions of generalization and automatic reward modeling) and knowledge-intensive reasoning (*e.g.,* question answering structured and unstructured data). Currently, I focus on recursive self-improvement in long-horizon reasoning environments.

Selected Publications

* equal contribution

[1] **X-Tree: Tokenizing Reusable Experience for Efficient Agent Generalization** [Paper] [Homepage]
Sitao Cheng, Xunjian Yin, Zhiyuan Sun, Yuxuan Li, Ruiwen Zhou, Xiangru Jian, Victor Zhong

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

[3] **Dual-Evolving Reinforcement Learning for Automated Reward Design** [Paper] [Code] [Homepage]
Sitao Cheng*, Tianle Li*, Xuhan Huang*, Xunjian Yin, Difan Zou

[4] **Understanding the Interplay between Parametric and Contextual Knowledge for Large Language Models** [Paper] [Code]
Sitao Cheng, Liangming Pan, Xunjian Yin, Xinyi Wang, William Yang Wang

[5] **Call Me When Necessary: LLMs can Efficiently and Faithfully Reason over Structured Environments** [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

[6] **QueryAgent: A Reliable and Efficient Reasoning Framework with Environmental Feedback-based Self-Correction** [Paper] [Code]
Xiang Huang*, **Sitao Cheng***, Shanshan Huang, Jiayu Shen, Yong Xu, Chaoyun Zhang, Yuzhong Qu

[7] **LEDOM: Reverse Language Model** [Paper] [Code]
Xunjian Yin, **Sitao Cheng**, Yuxi Xie, Xinyu Hu, Li Lin, Xinyi Wang, Liangming Pan, William Yang Wang, Xiaojun Wan

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

[9] **MarkQA: A Large Scale KBQA Dataset with Numerical Reasoning** [Paper] [Code]
Xiang Huang, **Sitao Cheng**, Yuheng Bao, Shanshan Huang, Yuzhong Qu

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

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

[12] **Question Decomposition Tree for Answering Complex Questions over Knowledge Bases** [Paper] [Code]
Xiang Huang, **Sitao Cheng**, Yiheng Shu, Yuheng Bao, Yuzhong Qu

Experience

Samaya AI Mountain View, USA
Research Intern · Advisor: Yuhao Zhang 08/2026 – Present

- **Recursive self-improvement of multi-turn reasoning agents.** Designing framework where an agent can explore, propose new tasks and learn in a loop without extra supervision in real-world complex reasoning environments.

University of Waterloo (R2L Lab) Waterloo, Canada
Ph.D. Student · Advisor: Prof. Victor Zhong 09/2025 – Present

- **Understanding the role of RL in compositional generalization.** Investigating whether RL acts as a synthesizer or an amplifier through controlled experiments, and studying the conditions for RL-driven generalization. Findings: RL genuinely synthesizes rather than amplifying probabilities, under the strict condition that the base model has acquired sufficient atomic skills. Results: One publication at ICLR 2026.
- **Dual-evolving reinforcement learning for automated reward design (DERL).** A dual-training framework in

which a reward proposer composes reward functions from task primitives and learns by policy gradient from the validation performance of the policies that its rewards train. RESULTS: One publication at ICLR 2026.

- **Hierarchy learning from limited experience.** Mining reusable structure from existing experience (trajectories, tasks with environmental feedback, etc.), inspired by byte-pair encoding. Integrating the hierarchy into training in three settings: offline RL, online RLVR, and on-policy self-distillation. RESULTS: One submission to ICLR 2027.

University of California, Santa Barbara (NLP Group) Santa Barbara, USA
Visiting Research Scholar · Advisor: Prof. William Wang 07/2024 – 06/2025

- **Understanding how effectively LLMs leverage parametric knowledge when contextual knowledge is given.** FINDINGS: LLMs consistently suppress their own knowledge when given context, regardless of the model, the knowledge type, the relation between the two knowledge sources, and the level of instruction. RESULTS: One publication at ACL 2025; four other collaborative projects published.

Microsoft Research Asia (DKI Group) Beijing, China
Research Intern & Mentor of Junior Interns · Advisors: Yong Xu, Fangkai Yang, Chaoyun Zhang 10/2023 – 06/2024

- **LLM reasoning frameworks over structured environments, via retrieval-augmented generation (Readi) or neural-symbolic reasoning (QueryAgent).** Explored how LLMs can reason both efficiently and faithfully over large-scale, heterogeneous structured environments, either by directly maintaining a reasoning path (Readi) or by building a query step by step (QueryAgent), each incorporating pertinent information for correction. RESULTS: Two publications at ACL 2024.
- **Efficient iterative retrieval with solely encoder-based models (EfficientRAG) and a new data organization paradigm (Thread) for RAG systems.** Modeled the links between text chunks by leveraging the understanding ability of LLMs, through new retrieval methods for smaller encoder-based models (EfficientRAG) and a reorganization of documents (Thread). RESULTS: One publication at EMNLP 2024 and one at EMNLP 2025.
- **LLM-based personalized AI assistant with structured knowledge graphs.** RESULTS: One submission to CHI.

Nanjing University (Websoft Lab) Nanjing, China
Student Researcher · Advisor: Prof. Yuzhong Qu 09/2021 – 06/2024

- **Step-by-step query building with self-correction from environmental feedback (QueryAgent).** Introduced a functional toolset with environmental feedback and a zero-shot correction method, for both reliability and efficiency. RESULTS: One publication at ACL 2024.
- **A KBQA benchmark requiring both multi-hop and numerical reasoning (MarkQA).** Proposed a structure requiring both capabilities and built a large-scale dataset (MarkQA) automatically. Designed a toolset with seamless SPARQL conversion as symbolic reasoning steps to reduce annotation burden. RESULTS: One publication at EMNLP 2023.
- **A question decomposition method for multi-hop reasoning over knowledge bases (QDT).** Designed a tree structure that represents complex questions, mapped to structured environments. RESULTS: One publication at AAAI 2023.

EDUCATION

University of Waterloo Waterloo, Canada
Ph.D. Student in Computer Science · Advisor: Prof. Victor Zhong 09/2025 – 06/2029 (expected)

Nanjing University Nanjing, China
M.S. in Computer Science and Technology · Advisor: Prof. Yuzhong Qu 09/2021 – 06/2024

University of Electronic Science and Technology of China Chengdu, China
B.E. in Software Engineering · Advisor: Prof. Yi Ding · GPA: 3.98 / 4.00 (Top 3) 09/2017 – 06/2021

HONORS AND AWARDS

Layer 6 – TD Graduate Scholarship in Data & AI, Fall 2025 University of Waterloo

MCM/ICM H Prize; Outstanding Student of Sichuan Province; Outstanding Student Award NJU, UESTC

OTHERS

Academic Service	ICML 2026 Silver Reviewer; ACL 2024 Volunteer; ARR, ICLR, NeurIPS Reviewer
Professional Skills	LLM applications, Transformers, attention mechanisms; PyTorch, Python, C++, SQL, LaTeX
Interests	Bodybuilding (SBD total over 6× body weight); basketball (department team); swimming
Social Service	Personal assistant to a senior impressionist artist at UC Santa Barbara