Zilong Wang

Ph.D. Student @ CSE, UC San Diego

✉ zlwang@ucsd.edu | ⌂ zilongwang.me

Research Interests

My research centers on leveraging language models for reasoning and planning, with a focus on:

- **LLM + Retrieval:** Enhancing factual accuracy and minimizing hallucinations through retrieval-augmented generation (RAG).
- **LLM + Code:** Improving the automatic debugging capabilities and coding proficiency of language models.
- **LLM + Multimodal:** Empowering language models to tackle multimodal tasks via autonomous agents, such as table-based reasoning and visually-rich document extraction.

Education

University of California, San Diego   *La Jolla, CA, USA*

Ph.D. in Computer Science and Engineering   *October 2020 - Now*

- Advised by Prof. Jingbo Shang

Peking University   *Beijing, CHN*

B.S. in Computer Science   *September 2016 - July 2020*

- Outstanding Graduate of Beijing City and Peking University (Top 5%)

Work Experience

Google Cloud AI Research & Google DeepMind   *Remote*

Student Researcher   *November 2023 - May 2024*

- Work with Chen-Yu Lee, Huaixiu Steven Zheng, Swaroop Mishra.
- Propose SPECULATIVE RAG to improve the effectiveness and efficiency of Retrieval Augmented Generation (RAG) in a map-reduce manner.
- SPECULATIVE RAG leverages a smaller specialist LM for drafting and a larger generalist LM for verification. This enables the RAG system to rapidly generate diverse drafts and effectively select the best one.

Google Cloud AI Research   *Sunnyvale, CA*

Student Researcher   *April 2023 - September 2023*

- Work with Chen-Yu Lee, Chun-Liang Li.
- Propose CHAIN-OF-TABLE to improve tabular reasoning ability of large language models. The paper is accepted by ICLR '24.
- The algorithm has been integrated into the products in Google Cloud.
- CHAIN-OF-TABLE enables step-by-step reasoning with table inputs by simplifying complex tables into more informative and manageable segments. LLMs can now reason about tables more effectively similar to the thought process of a human analyst.

Google DeepMind   *Mountain View, CA*

Student Researcher   *June 2022 - September 2022*

- Work with Sandeep Tata, Yichao Zhou, Wei Wei, Chen-Yu Lee.
- Propose a new public benchmark for visually-rich document understanding focusing on the few-shot setting and template structure learning. The paper is accepted by KDD '23.
- Propose A type-aware evaluation toolkit to evaluate the performance of language models in visually-rich document understanding.

Adobe Research   *Remote*

Research Intern   *June 2021 - September 2021*

- Work with Vlad Morariu, Tong Sun.
- Propose a layout-aware pre-trained language model for visually-rich document understanding using multi-granular inputs to encode template structure. The paper is accepted by EMNLP '22.

Microsoft Research Asia   *Beijing, CHN*

Research Intern   *September 2020 - March 2021*

- Work with Lei Cui.
- Propose an auto-regressive language model to encode multi-modal features and extract reading order in visually-rich documents. The paper is accepted by EMNLP '21.
- Propose an automatic approach to build large reading order dataset from XML metadata of Microsoft Office files.

Papers & Pre-prints

[Google Scholar] [DBLP]

1. OfficeBench: Benchmarking Language Agents across Multiple Applications for Office Automation
   Zilong Wang, Yuedong Cui, Li Zhong, Zimin Zhang, Da Yin, Bill Yuchen Lin, Jingbo Shang
   (Preprint)

2. Speculative RAG: Enhancing Retrieval Augmented Generation through Drafting
   Zilong Wang, Zifeng Wang, Long Le, Huaixiu Steven Zheng, Swaroop Mishra, Vincent Perot, Yuwei Zhang, Anush Mattapalli, Ankur Taly, Jingbo Shang, Chen-Yu Lee, Tomas Pfister
   (Preprint)

3. Chain-of-Table: Evolving Tables in the Reasoning Chain for Table Understanding
   Zilong Wang, Hao Zhang, Chun-Liang Li, Julian Martin Eisenschlos, Vincent Perot, Zifeng Wang, Lesly Miculicich, Yasuhisa Fujii, Jingbo Shang, Chen-Yu Lee, Tomas Pfister
   The International Conference on Learning Representations (**ICLR**), 2024

4. LDB: A Large Language Model Debugger via Verifying Runtime Execution Step-by-step
   Li Zhong, **Zilong Wang**, Jingbo Shang
   Findings of the Association for Computational Linguistics: ACL 2024 (**ACL**), 2024

5. Answer is All You Need: Instruction-following Text Embedding via Answering the Question
   Letian Peng, **Zilong Wang**, Feng Yao, Zihan Wang, Jingbo Shang
   Annual Meeting of the Association for Computational Linguistics (**ACL**), 2024

6. LMDX: Language Model-based Document Information Extraction and Localization
   Vincent Perot, Kai Kang, Florian Luisier, Guolong Su, Xiaoyu Sun, Ramya Sree Boppana, **Zilong Wang**, Jiaqi Mu, Hao Zhang, Nan Hua
   Findings of the Association for Computational Linguistics: ACL 2024 (**ACL**), 2024

7. Can ChatGPT replace StackOverflow? A Study on Robustness and Reliability of Large Language Model Code Generation
   Li Zhong, **Zilong Wang**
   The 38th Annual AAAI Conference on Artificial Intelligence (**AAAI**), 2023

8. Towards Zero-shot Relation Extraction in Web Mining: A Multimodal Approach with Relative XML Path
   Zilong Wang, Jingbo Shang
   Findings of the Association for Computational Linguistics: EMNLP 2023 (**EMNLP**), 2023

9. VRDU: A Benchmark for Visually-rich Document Understanding
   Zilong Wang, Yichao Zhou, Wei Wei, Chen-Yu Lee, Sandeep Tata
   Proc. of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**), 2023

10. MGDoc: Pre-training with Multi-granular Hierarchy for Document Image Understanding
    Zilong Wang, Jiuxiang Gu, Chris Tensmeyer, Nikolaos Barmpalios, Ani Nenkova, Tong Sun, Jingbo Shang and Vlad I. Morariu
    Proc. of the 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2022

11. Formulating Few-shot Fine-tuning Towards Language Model Pre-training: A Pilot Study on Named Entity Recognition
    Zihan Wang, Kewen Zhao, **Zilong Wang**, Jingbo Shang
    Findings of the Association for Computational Linguistics: EMNLP 2022 (**EMNLP**), 2022

12. Towards Few-shot Entity Recognition in Document Images: A Label-aware Sequence-to-Sequence Framework
    Zilong Wang, Jingbo Shang
    Findings of the Association for Computational Linguistics: ACL 2022 (**ACL**), 2022

13. LayoutReader: Pre-training of Text and Layout for Reading Order Detection
    Zilong Wang, Yiheng Xu, Lei Cui, Jingbo Shang, Furu Wei
    Proc. of the 2021 Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021

14. DocStruct: A Multimodal Method to Extract Hierarchy Structure in Document for General Form Understanding
    Zilong Wang, Mingjie Zhan, Xuebo Liu and Ding Liang
    Findings of the Association for Computational Linguistics: EMNLP 2020 (**EMNLP**), 2020

15. Exploring Semantic Capacity of Terms
    Jie Huang*, **Zilong Wang***, Kevin Chang, Wen-mei Hwu and JinJun Xiong
    Proc. of the 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2020

Honors & Awards

2020-2021	UCSD Powell Fellowship
2020-2021	UCSD Jacob School of Engineering Fellowship
2020	Summa Cum Laude of Beijing City
2020	Summa Cum Laude of Peking University

Professional Service

Reviewer   ICLR (2024), NeurIPS (2022, 2023), ACL (2022, 2023) EMNLP (2022, 2023), ICML (2024), COLM (2024), CVPR (2024), CtrlGen Workshop(2021), COLING (2022)

Skills

Skills   Machine Learning, Natural Language Processing, Language Modeling, Multimodal Data Mining, Visually-rich Document Understanding, Webpage Information Mining

Programming   Python (PyTorch, Tensorflow, Keras), Java, C++, C#, Javascript

References available upon request.