Zilong (Ryan) Wang

La Jolla, CA   zlwang@ucsd.edu   zlwang-cs.github.io

Research Interests

My research spans several areas of natural language processing, including reasoning, information extraction, multimodal learning, and language modeling. I primarily focus on LLM post-training, targeting at aligning LLMs to complex reasoning and planning with knowledge-intensive queries, autonomous agents, and mathematical problem-solving. Currently, I am exploring the collaboration of weak and strong LLMs and the process supervision with Monte Carlo Tree Search.

Education

University of California, San Diego   *Oct. 2020 – Now*
Ph.D. in Computer Science

- Advised by Prof. Jingbo Shang

Peking University   *Sept. 2016 – Jun. 2020*
B.S. in Computer Science

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

Experience

Amazon, Nile Foundation Team   *Palo Alto, CA*
Research Scientist Intern   *Jun. 2024 - Now*

- Working on efficient scaling up of process supervision on multi-step problem solving (e.g. agents, math reasoning).

Google Cloud AI Research & Google DeepMind   *Remote*
Student Researcher   *Nov. 2023 - May 2024*

- Working on reducing hallucinations and enhancing efficiency for retrieval augmented generation (RAG).

Google Cloud AI Research   *Sunnyvale, CA*
Student Researcher   *Apr. 2023 - Sept. 2023*

- Working on improving tabular reasoning ability of large language models through autonomous agents.
- *The Chain-of-Table algorithm has been integrated into the products in Google Cloud.*

Google DeepMind   *Mountain View, CA*
Student Researcher   *Jun. 2022 - Sept. 2022*

- Working on multimodal information extraction and visually-rich document understanding (e.g. forms, receipts, structured documents).

Adobe Research, Document Intelligence Lab   *Remote*
Research Intern   *Jun. 2021 - Sept. 2021*

- Working on enabling language models to understand visual features from structured documents (e.g. forms, receipts, structured documents).

Publications

1. **Rising Rewards Elite Efficient Process Supervision in Automated Agent Training**
   Zilong Wang, Jingbo Shang, Amazon Nile Foundation Model Team.
   (ongoing)

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. **TableRAG: Million-Token Table Understanding with Language Models**

Si-An Chen, Lesly Miculicich, Julian Martin Eisenschlos, Zifeng Wang, **_Zilong Wang_**, Yanfei Chen, Yasuhisa Fujii, Hsuan-Tien Lin, Chen-Yu Lee, Tomas Pfister
Conference on Neural Information Processing Systems (**NeurIPS**), 2024

4. **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

5. **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)

6. **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

7. **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

8. **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

9. **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

10. **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

11. **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

12. **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

13. **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

14. **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

15. **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

16. **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

17. **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 Peking University (Top 5%)

Professional Service & Skills

Reviewer: ICLR, NeurIPS, ACL, EMNLP, NAACL, CVPR, ICML, COLM, COLING

Skills: Large language model reasoning and planning, reinforcement learning, DPO, PPO, multimodal data mining