

Wenyue Hua

PostDoc in Computer Science, University of California, Santa Barbara ◇ Email: wenyue.hua@rutgers.edu

EDUCATION

Rutgers University, New Brunswick

09/2020 — 10/2024

Ph.D. in Computer Science

GPA: 3.93/4.0

Thesis title: Trustworthy Large Language Models; Advisor: Professor Yongfeng Zhang

Rutgers University, New Brunswick

09/2018 — 10/2020

Master in Arts , Linguistics

• GPA: 4.0/4.0

Thesis title: Learning Underlying Representations and Phonological Grammars; Advisor: Professor Adam Jardine

University of California, Los Angeles (UCLA)

09/2014 — 06/2018

BS in Mathematics, General & BA in Linguistics&Philosophy with Specialization in Computing

• GPA: 3.75/4.0

SELECTED PUBLICATIONS

- [1] (submitted to ICLR) **Wenyue Hua**, Mengting Wan, Shashank Vadrevu, Ryan Nadel, Yongfeng Zhang, Chi Wang. Interactive Speculative Planning: Enhance Agent Efficiency through Co-design of System and User Interface
- [2] (submitted to NAACL) **Wenyue Hua**, Kaijie Zhu, Lingyao Li, Lizhou Fan, Shuhang Lin, Mingyu Jin, Haochen Xue, Zelong Li, Jindong Wang, Yongfeng Zhang. Disentangling Logic: The Role of Context in Large Language Model Reasoning Capabilities
- [3] Shuhang Lin*, **Wenyue Hua***, Hang Hua, Jianchao Ji, Lizhou Fan, Lingyao Li, Jiebo Luo, Yongfeng Zhang *BattleAgent: Multi-modal Dynamic Emulation on Historical Battles to Complement Historical Analysis* The 2024 Conference on Empirical Methods in Natural Language Processing, demo track
- [4] **Wenyue Hua**, Xianjun Yang, Mingyu Jin, Zelong Li, Wei Cheng, Yongfeng Zhang *TrustAgent: Towards Safe and Trustworthy LLM-based Agents through Agent Constitution* The 2024 Conference on Empirical Methods in Natural Language Processing
- [5] Lizhou Fan*, **Wenyue Hua***, Lingyao Li, Haoyang Ling, Yongfeng Zhang *NPHardEval: Benchmarking Reasoning Ability of Large Language Models via Complexity Classes* The 62nd Annual Meeting of the Association for Computational Linguistics
- [6] **Wenyue Hua***, Jiang Guo*, Mingwen Dong, Henghui Zhu, Patrick Ng, Zhiguo Wang. *Propagation and Pitfalls: Exploring the Challenges of Knowledge Editing through Counterfactual Tasks*. The 62nd Annual Meeting of the Association for Computational Linguistics
- [7] (submitted for ICLR) **Wenyue Hua***, Lizhou Fan*, Lingyao Li, Kai Mei, Jianchao Ji, Yingqiang Ge, Libby Hemphill, Yongfeng Zhang. *War and Peace (WarAgent): Large Language Model-based Multi-Agent Simulation of World Wars*.
- [8] Yingqiang Ge, **Wenyue Hua**, Jianchao Ji, Juntao Tan, Shuyuan Xu, Yongfeng Zhang. *OpenAGI: When LLM Meets Domain Experts..* 37th Conference on Neural Information Processing Systems.
- [9] **Wenyue Hua**, Shuyuan Xu, Yingqiang Ge, Yongfeng Zhang. *How to Index Item IDs for Recommendation Foundation Models*. 1st International ACM SIGIR Conference on Information Retrieval in the Asia Pacific.
- [10] **Wenyue Hua**, Yingqiang Ge, Shuyuan Xu, Jianchao Ji, Zelong Li, Yongfeng Zhang. *Towards Fairness-aware Large Foundation Models for Recommendation based on Counterfactually-fair Prompting*. 18th Conference of the European Chapter of the Association for Computational Linguistics
- [11] Wenzheng Zhang, **Wenyue Hua**, Karl Stratos. *EntQA: Entity Linking as Question Answering*. International Conference of Learning Representations 2022. Spotlight presentation.

[12] **Wenyue Hua**, Adam Jardine. *Learning Input Strictly Local Functions From Their Composition*. Proceedings of Machine Learning Research. (highest scored submission)

INDUSTRY EXPERIENCES

Efficient Agent Planning with System Design and Human-AI Collaboration

Research intern, AutoGen, Microsoft Research

05/2024 — 08/2024

- Design the first algorithm that accelerates agent planning: speculative planning, which reduces latency of agent planning for 40%
- Design a scheduling mechanism to enable an easy-to-follow user interface on top of speculative planning
- Design an algorithm to accept active user interaction without affecting latency acceleration

Knowledge Propagation in Large Language Model based on Model Editing

Applied Scientist, AWS AI Lab, Quicksight-Q, Amazon

05/2023 — 08/2023

- Design a knowledge editing benchmark containing seven different types of discrete reasoning to study whether a model can reason based on newly learned knowledge
- Evaluate and compare different model editing methods on the benchmark, including input-augmentation, QLoRA tuning, MEMIT, and MEND.
- study the effects of (1) model sizes (2) reasoning types (3) editing methods

Automatic Error Detection model in Classification tasks of Natural Language Processing

Research Scientist, AI Lab, Tencent America

05/2022 — 08/2022

- Design evaluation metric: rejection experiment, synthetic data recovery test, active learning
- Design a comprehensive list of features and statistic significance tests to explain error discovered by the model
- Develop neural network based Automatic Error Detection models and develop the model from a simple binary classification task to multi-label classification task and sequence to sequence task

SELECTED AWARDS

- National Science Foundation, SBIR (\$50,000) 08/2021
- SGS Travel Grant (\$175) 01/2020
- College Honor 06/2018
- Phi Beta Kappa 06/2018
- Department Honor 06/2018
- Dean's Honor List 06/2014 — 09/2018

SELECTED EXTRA-CURRICULAR ACTIVITIES

Graduate Student Liaison to Office of Advanced Research Computing

09/2019 — 09/2020

Rutgers University

- Act as a point of contact for the department graduate student users of Amarel for resources and trainings.
- Add to department's documentation about Amarel as an augmentation of Amarel docs for specific disciplines.

President

06/2017 — 06/2018

Driftwood Seminar, UCLA

- Designed and held three speaker panels with five to six speakers for approximately fifty participants across various social, cultural and economic topics.
- Organized and coordinated the work among academic, marketing and operation teams.

Academic Director

06/2016 — 06/2017

Driftwood Asian-Pacific Seminar, UCLA

- Designed topics and content of each panel and talk.
- Co-founder of the club of Driftwood Asian-Pacific Seminar.

SKILLS

Languages	Chinese (proficient), English (proficient)
Programming Languages	Python, C++, Haskell, MATLAB, R
Computer Skills	Git, L ^A T _E X, Linux, Pytorch, MS Offices