Yi Wu

Email: mail.yi.wu@gmail.com Page: https://page-yiwu.github.io Mobile: +1 309 703 6663

RESEARCH INTERESTS

My current research aims on exploring and applying LLM-based agentic systems into:

- Unifying cognitive concepts in the architecture of agents on strategic reasoning games like StarCraft II.
- Grounded persuasive agents for real-estate marketing with user preference alignment.
- Counterexample-guided proof repair to automate Rust program verification.

EDUCATION

- **The University of Chicago** Chicago, IL
  Pre-Doctoral M.S. in Computer Science *2024.9-2025.12 (expected)*
  Faculty collaborators: Prof. Kexin Pei, Prof. Haifeng Xu
- **University of Wisconsin-Madison** Madison, WI
  B.S. in Computer Sciences *2022.9-2024.8*
  GPA: 4.0/4.0
- **University of Illinois Urbana-Champaign** Champaign, IL
  The Grainer College of Engineering, Exchange Program *2022.1-2022.5*
  GPA: 3.91/4.0
- **The Hong Kong University of Science and Technology** Hong Kong SAR
  B.S. in Computer Science, Mathematics *2019.9-2022.8*
  GPA: 3.716/4.3

Related Courses:
Honors Design and Analysis of Algorithms, Honors Probability, Honors Linear Algebra, Calculus I, II, III, Math Toolkit, Communication Networks, Intro-Artificl Intelligence, Deep Learning in Computer Vision, Advanced Natural Language Processing, Intro to Optimization, Intro to Machine Perception

PUBLICATIONS & PREPRINTS

[1] **AI Realtor: Towards Grounded Persuasive Language Generation for Automated Copywriting**
Chenghao Yang*, Jibang Wu*, **Yi Wu**, Simon Mahns, Chaoqi Wang, Hao Zhu, Fei Fang, Haifeng Xu
In submission. [preprint]

[2] **LLMs Aren't Good Strategists, Yet Can Accumulate Episodes for Improved Planning**
Yi Wu*, Zhimin Hu*
In *Reasoning and Planning for LLMs, ICLR 2025*. [pdf]

[3] **Evolving Domain Adaptation of Pretrained Language Models for Text Classification**
Yun-Shiuan Chuang, **Yi Wu**, Dhruv Gupta, Rheeya Uppaal, Ananya Kumar, Luhang Sun, Makesh M. Sreedhar, Sijia Yang, Timothy T. Rogers, Junjie Hu
In *NeurIPS 2023 Workshop on Distribution Shifts (DistShift)*.[pdf]; Longer version [pdf]

[4] **KnowComp Submission for WMT23 Word-Level AutoCompletion Task**
Yi Wu, Haochen Shi, Weiqi Wang, Yangqiu Song
In *Proceedings of the Eighth Conference on Machine Translation (WMT-2023), EMNLP*. [pdf], [code]

RESEARCH EXPERIENCES

Sec Lab, University of Chicago *Advisor: Kexin Pei* *2025.5 - present*

- **ExVerus: Counterexample-Guided Proof Repair for Rust (Verus) Programs**
  - Co-developed *ExVerus*, an agentic verification framework that reframes invariant repair as an inductive reasoning task guided by calling tools to generate concrete counterexamples.
  - Designed modules for counterexample generation, validation, and generalization, enabling LLMs to block counterexamples rather than rely on weak abductive reasoning.
  - Demonstrated significant gains over state-of-the-art systems like AutoVerus on VerusBench and new benchmarks, reducing token cost while improving proof success rates.

SIGMA Lab, University of Chicago *Advisor: Haifeng Xu* *2025.4 - present*

- **AI Realtor: Grounded Persuasive Copywriting Agents**
  - Helped to build an agentic workflow with three modules: grounding (attribute-to-feature mapping), personalization (eliciting buyer preferences), and marketing (capturing localized surprising features).
  - Led the design of fine-grained hallucination checks by distinguishing hard vs. soft factual attributes, creating quantitative faithfulness scores, and demonstrating that AI Realtor surpasses both human-written and fine-tuned baselines in factual accuracy.
  - Achieved a 70% win rate over professional human-written listings in pairwise comparisons, while maintaining factual accuracy through rigorous hallucination checks

Independent Research Project: EpicStar *2024.12 - 2025.4*

- **LLMs Aren't Good Strategists, Yet Memory-Enhanced Agency Boosts Reasoning**
  - Developed the *EpicStar* framework, augmenting LLM agents with episodic and working memory to enable long-term strategic planning and adaptive short-term reasoning in StarCraft II.
  - Ran large-scale experiments showing EpicStar achieves higher win rates with significantly fewer tokens compared to vanilla LLM baselines .
  - Identified fundamental gaps between pattern-matching and true strategic reasoning, offering insights for designing memory-augmented agents in sequential decision-making tasks.

Hulab & Knowledge and Concepts Lab, UW-Madison *2023.3 - 2023.12*
Advisors: Junjie Hu, Timothy T. Rogers

- **Simulating Opinion Dynamics with Networks of LLM-based Agents**
  - Helped to explore the use of Large Language Models (LLMs) for simulating human opinion in group dynamics in politically charged environments.
  - Proposed ideas to identify biases in LLM agents towards accurate information, impacting the simulation of resistant viewpoints, like in climate change debates.
  - Helped to utilize LLMs to role-play partisan personas, finding that responses without Chain-of-Thought (CoT) reasoning align more with human behaviors.
- **Evolving Domain Adaptation of Pretrained Language Models for Text Classification**
  - Investigated Evolving Domain Adaptation (EDA) strategies for Pretrained Language Models (PLM) in time-series text classification, especially incremental self-training.
  - Conducted extensive experiments to demonstrate incremental self-training's superiority in adapting PLMs to evolving domain shifts.
  - Suggested the necessity of regular PLM updates for sustained real-world application accuracy and suggested future research on PLM robustness to natural language evolution.

KnowComp, HKUST *Advisor: Yangqiu Song* *2023.3 - 2023.8*

- **WMT23 Word-Level AutoCompletion Task**
  - Proposed a LLM-based system for the WMT23 Word-Level Auto-Completion (WLAC) task, using LLMs to evaluate performance in multilingual contexts.
  - Tested the system in Chinese-English, German-English, and English-German translation directions.
  - Assessed performance under zero-shot and few-shot settings, finding improved accuracy with additional training exemplars.

HONORS & AWARDS

- **Dean's List**, 2019, HKUST
- **Dean's List**, 2022 Spring, UIUC
- **Dean's List**, 2022 Fall, 2023 Spring, 2023 Fall, 2024 Spring, UW-Madison
- **University's Scholarships Scheme for Continuing Undergraduates**, 2020-2021, 2021-2022, HKUST
- **Merit Based Scholarship**, 2024-2025, University of Chicago

SKILLS

- **Programming/Tools**: Python, C/C++, Java, Rust, SQL; PyTorch, Hugging Face, Git, LaTeX;
- **Languages**: Chinese (Native), English (Professional)