HONGRU WANG

Emails: hrwang@se.cuhk.edu.hk / hongru.carrywang@gmail.com

Homepage | Thesis Slides (v1) | Google Scholar

RESEARCH STATEMENT

My research focus revolves around reasoning and acting of personalized language agents, designed to seamlessly unifing them from tool perspective by regarding reasoning as internal cognitive tools (ollike reasoning) while acting as external physical tools instead of treat them in isolation. My long-term objective is to achieve the *impossible triangle* between safety, personalization and autonomy of language agent.

EDUCATION

The Chinese University of Hong Kong2021/09 - NowPh.D. candidate supervised by Prof. Kam-Fai WongGPA: 3.82/4.00

Dialogue System, Large Language Models, Tool Learning

The University of Edinburgh 2024/04 - Now

Visiting Postgraduate student supervised by Prof. Jeff Z. Pan Dialogue System, Large Language Models, Tool Learning

The Chinese University of Hong Kong M.Eng. Computer Science and Engineering GPA: 3.60/4.00

Data Mining, Big Data Technology, Machine Learning

Communication University of China 2015/09 - 2019/09

B.Eng. Computer Science and Technology GPA: 3.27/4.00

Computer Operation System, Computer Networks, Data Structures and Algorithms

(CO-)FIRST AUTHOR CONFERENCES

 AppBench: Planning of Multiple APIs from Various APPs for Complex User Instruction Hongru Wang, Rui Wang, Boyang XUE, et. al, Jeff Z. Pan, Kam-Fai Wong EMNLP 2024 LLM, Tool Learning (Apple Intelligence)

9. TPE: Towards Compositional Reasoning over Conceptual Tools with Multi-persona Collaboration Hongru Wang, Huimin Wang, Lingzhi Wang, et. al, Kam-Fai Wong

NLPCC 2024 LLM , Tool Learning (Meta-reasoning theory / Cognitive tools)

- 8. Enhancing Large Language Models Against Inductive Instructions with Dual-critique Prompting Rui Wang*, Hongru Wang*, Fei Mi, Boyang XUE, Yi Chen, Kam-Fai Wong, Ruifeng Xu

 NAACL 2024 LLM, Safety * denotes equal contribution
- 7. UniRetriever: Multi-task Candidates Selection for Various Context-Adaptive Conversational Retrieval Hongru Wang, Boyang Xue, Baohang Zhou, et. al, Kam-Fai Wong

 LREC-COLING 2024 Dialogue System, Tool Learning
- M3Sum: A Novel Unsupervised Language-guided Video Summarization
 Hongru Wang, Baohang Zhou, Zhengkun Zhang, Yiming Du, David Ho, Kam-Fai Wong
 ICASSP 2024 LLM, MM
- 5. Large Language Models as Source Planner for Personalized Knowledge-grounded Dialogue Hongru Wang, Minda Hu, Yang Deng, et. al, Irwin King, Kam-Fai Wong Findings of EMNLP 2023 Dialogue System, LLM Best Paper Award @ @Doctoral Forum

- 4. Cue-CoT: Chain-of-thought Prompting for Responding to In-depth Dialogue Questions with LLMs Hongru Wang*, Rui Wang*, Fei Mi, et. al, Ruifeng Xu and Kam-Fai Wong
 Findings of EMNLP 2023 Dialogue System, LLM Blog
- 3. MCML: A Novel Memory-based Contrastive Meta-Learning Method for Few Shot Slot Tagging Hongru Wang, Zezhong Wang, Wai-Chung Kwan, Kam-Fai Wong IJCNLP-AACL 2023 Dialogue System
- Integrating Pretrained Language Model for Dialogue Policy Learning <u>Hongru Wang</u>, Huimin Wang, Zezhong Wang and Kam-Fai Wong
 <u>ICASSP 2022</u> Dialogue System , RLHF (probably first work of RLHF)
- CUHK at SemEval-2020 Task 4: CommonSense Explanation, Reasoning and Prediction ..
 Hongru Wang, Xiangru Tang, Sunny Lai, et. al, Gabriel Pui Cheong Fung and Kam-Fai Wong SemEval of COLING 2020. Commonsense

(CO-)FIRST AUTHOR JOURNAL

- 2. KddRES: A Multi-level Knowledge-driven Dialogue .. Towards Customized Dialogue System Hongru Wang, Wai-Chung Kwan, Min Li, Zimo Zhou and Kam-Fai Wong Computer Speech and Language 2024 (SCI Q2, IF: 4.3) Dialogue System
- A Survey on Recent Advances and Challenges in Reinforcement Learning Methods ..
 Wai-Chung Kwan*, <u>Hongru Wang*</u>, Huimin Wang and Kam-Fai Wong
 Machine Intelligence Research 2023 Dialogue System , RLAIF

NON-FIRST AUTHOR CONFERENCES

- 17. AutoCV: Empowering Reasoning with Automated Process Labeling via Confidence Variation Jianqiao Lu, Zhiyang Dou, <u>Hongru Wang</u>, et. al, Zhijiang Guo

 NeruIPS 2024 LLM (Meta-reasoning theory)
- Knowledge Conflicts for LLMs: A Survey Rongwu Xu, Zehan Qi, Zhijiang Guo, Cunxiang Wang, Hongru Wang, Yue Zhang, Wei Xu EMNLP 2024 LLM Blog
- 15. VLEU: a Method for Automatic Evaluation for Generalizability of Text-to-Image Models Jingtao Cao, Zhang Zheng, Hongru Wang, Kam-Fai Wong

 EMNLP 2024 MM
- Less is More: Making Smaller Language Models .. Subgraph Retrievers for Multi-hop KGQA Wenyu Huang, Guancheng Zhou, <u>Hongru Wang</u>, et. al, Mirella Lapata, Jeff Z. Pan Findings of EMNLP 2024 LLM (Differentiable Search Index (DSI))
- 13. Enhancing Biomedical Knowledge RAG with Self-Rewarding Tree Search and PPO Minda Hu, Licheng Zong, Hongru Wang, et. al, Kam-Fai Wong, Yu Li, Irwin King Findings of EMNLP 2024 LLM
- 12. DPDLLM: A Black-box Framework for Detecting Pre-training Data from Large Language Models Baohang Zhou, Zezhong WANG, Lingzhi Wang, Hongru Wang, et. al, Kam-Fai Wong Findings of ACL 2024 LLM
- 11. Medical Dialogue: A Survey of Categories, Methods, Evaluation and Challenges Xiaoming Shi, Zeming Liu, Li Du, Yuxuan Wang, Hongru Wang, et. al, Shaoting Zhang Findings of ACL 2024 LLM

- 10. REGA: Role Prompting Guided Multi-Domain Adaptation for Large Language Models Rui Wang, Fei Mi, Yi Chen, Boyang XUE, Hongru Wang, Qi Zhu, Kam-Fai Wong, Ruifeng Xu Findings of **NAACL 2024** LLM, Role Playing
- 9. SELF-GUARD: Empower the LLM to Safeguard Itself Zezhong Wang, Fangkai Yang, Lu Wang, Pu Zhao, Hongru Wang, et. al, Kam-Fai Wong NAACL 2024 LLM, Safety
- 8. JoTR: A Joint Transformer and Reinforcement Learning Framework for Dialog Policy Learning Wai-Chung Kwan, Huimin Wang, Hongru Wang, et. al, Kam-Fai Wong LREC-COLING 2024 Dialogue System
- 7. MCIL: Multimodal Counterfactual Instance Learning for .. Multimodal Information Extraction Baohang Zhou, Ying Zhang, Kehui Song, Hongru Wang, Yu Zhao, Xuhui Sui, Xiaojie Yuan LREC-COLING 2024 LLM, MM
- ReadPrompt: A Readable Prompting Method for Reliable Knowledge Probing Zezhong Wang*, Luyao YE*, <u>Hongru Wang</u>, Wai-Chung Kwan, David Ho, Kam-Fai Wong Findings of EMNLP 2023 LLM
- 5. Improving Factual Consistency for Knowledge-Grounded Dialogue Systems via Knowledge .. Boyang XUE*, Weichao Wang*, Hongru Wang, et. al, Xin Jiang, Qun Liu, Kam-Fai Wong Findings of EMNLP 2023 Dialogue System
- 4. Prompting and Evaluating Large Language Models for Proactive Dialogues .. Yang Deng, Lizi Liao, Liang CHEN, Hongru Wang, Wenqiang Lei, Tat-Seng Chua Findings of EMNLP 2023 LLM
- 3. Towards Robust Personalized Dialogue Generation via Order-Insensitive Representation .. Liang Chen, <u>Hongru Wang</u>, Yang Deng, Wai Chung Kwan, Zezhong Wang, Kam-Fai Wong Findings of ACL 2023 (short) Dialogue System
- Retrieval-free Knowledge Injection through Multi-Document Traversal for Dialogue Models Rui Wang, Jianzhu Bao, Fei Mi, Yi Chen, Hongru Wang, et. al, Kam-Fai Wong, Ruifeng Xu ACL 2023 Dialogue System
- DIGAT: Modeling News Recommendation with Dual-Graph Interaction Zhiming Mao, Jian Li, Hongru Wang, Xingshan Zeng, Kam-Fai Wong Findings of EMNLP 2022 Others

WORKSHOP, TUTORIAL AND OTHERS

- 3. PerLTQA: A Personal Long-Term Memory Dataset for Memory Classification, Retrieval, and .. Yiming Du, Hongru Wang, Zhengyi Zhao, et. al, Kam-Fai Wong SIGHAN Workshop of ACL 2024. Best Paper Award ②.
- 2. OSPC: Detecting Harmful Memes with Large Language Model as a Catalyst Jingtao Cao, Zheng Zhang, Hongru Wang, Bin Liang, Hao Wang, Kam-fai Wong Online Safety Prize Challenge, WWW 2024. Champion Solution ©.
- Empowering Large Language Models: Tool Learning for Real-World Interaction Hongru Wang, Yujia Qin, Yankai Lin, Jeff Z. Pan, Kam-fai Wong Tutorial, SIGIR 2024 LLM, Tool Learning

PREPRINTS

- 4. Can LLMs Evaluate Complex Attribution in QA? Automatic Benchmarking Using Knowledge Graphs Nan Hu, Jiaoyan Chen, Yike Wu, Guilin Qi, Hongru Wang, Sheng Bi, Tongtong Wu, Jeff Z. Pan Under Review, ICLR 2025 LLM
- 3. A Survey of the Evolution of Language Model-Based Dialogue Systems

 Hongru Wang, Lingzhi Wang, Yiming Du, Liang Chen, Jingyan Zhou, Yufei Wang, Kam-Fai Wong

 Under Review LLM Dialogue System Blog
- 2. Self-DC: When to retrieve and When to generate? Self Divide-and-Conquer for Compositional .. Hongru Wang*, Boyang Xue*, Baohang Zhou, et. al, Kam-Fai Wong

 Under Review LLM, Tool Learning
- UniMS-RAG: A Unified Multi-source Retrieval-Augmented Generation for Personalized ...
 Hongru Wang, Wenyu Huang, Yang Deng, et. al, Jeff Z Pan, Kam-Fai Wong
 Under Review LLM , Dialogue System

EXPERIMENCE

Research Intern Seed-LLM-Harizon (Doubao) 2024/08 - Now ByteDance (SZ) (Mentor: Dr. Deng Cai, Dr. Wanjun Zhong)

Research Intern Speech and Semantic, Noah's Ark Lab 2022/04 - 2023/03 Huawei (SZ) (Mentor: Dr. Fei Mi, Dr. Weichao Wang, Prof. Qun Liu)

Research Assistant MoE Key Lab of High Confidence Software Technologies **2019/12 - 2021/07 The Chinese University of Hong Kong** (Mentor: Prof. Kam-Fai Wong)

GRANTS / FUNDING

- 3. Overseas Research Attachment Programme (50,000 HKD, ORAP 2023-2024)
- 2. A Knowledge Graph Based Dynamic Video Extractive Summarization System (PRP/054/21FX)
- 1. "SEVES: Semantic-driven Effective Video Extractive Summarization System" Technology and Business Development Fund (200,000 HKD, TBF22ENG004)

TEACHING ASSISTANT

SEEM 3450 Engineering Innovation and Entrepreneurship SEEM 3490 Information Systems Management SEEM 5730 / ECLT 5910 Information Technology Management

COMPETITIONS & AWARDS

Top 1 at Online Safety Prize Challange, WWW 2024	international
Top 10 at ICLR 2021 Workshop MLPCP Track 1 (rank 7th)	international
Third Price at SMP2020-ECDT Few-shot Spoken Language Understanding	international
Top 10 at SemEval2020-Task4: CommonSense Detection and Explanations	international
Distinguished Academic Performance Scholarship (2019-2020), CUHK CSE	school
Meritorious Winner, Mathematical Contest In Modeling (2018)	international
A software copyright of "WeCampus WeChat mini-program" (2018SR562540)	national
Third Prize of Internet + College Students Innovation and Entrepreneurship Compe	tition school
China National Radio Scholarship	school