

# Yi CHENG

✉ alyssa.cheng@connect.polyu.hk    ☎ (+852) 54866325

🌐 Homepage: <https://yicheng98.github.io/>

## EDUCATION

**The Hong Kong Polytechnic University**

**Hong Kong, China**

*Ph.D. in Computer Science*

Sep 2021 – Present

◦ **Advisor:** Prof. Maggie Wenjie Li

◦ **Research Interests:** Conversational AI, Emotional Support Conversations, Medical Consultation Dialogues

**Peking University**

**Beijing, China**

*B.S. in Computer Science*

Sep 2016– June 2020

◦ **Advisor:** Associate Prof. Sujian Li

**University of Oxford**

**Oxford, United Kingdom**

*Summer Program*

July 2018 – Aug 2018

## PUBLICATIONS

(\*indicates equal contribution)

- **Evolving to be your soulmate: Towards Self-evolving Conversational Agents for Companionship with Dynamically adapted persona.**

Yi Cheng, Wenge Liu, Chak Tou Leong, Xiao Yang, Yi Ouyang, Wenjie Li, Xian Wu, and Yefeng Zheng.

*Submitted to The 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024 Submission).*

- **COOPER: Coordinating Specialized Agents towards a Complex Dialogue Goal.**

Yi Cheng, Wenge Liu, Jian Wang, Chak Tou Leong, Yi Ouyang, Wenjie Li, Xian Wu, and Yefeng Zheng.

*In Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI 2023).*

- **Self-Detoxifying Language Models via Toxification Reversal.**

Chak Tou Leong\*, Yi Cheng\*, Jiashuo Wang, Jian Wang, and Wenjie Li.

*In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2023).*

- **Improving Multi-turn Emotional Support Dialogue Generation with Lookahead Strategy Planning.**

Yi Cheng\*, Wenge Liu\*, Wenjie Li, Jiashuo Wang, Ruihui Zhao, Bang Liu, Xiaodan Liang, Yefeng Zheng.

*In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP'2022).*

- **MedDG: An Entity-Centric Medical Consultation Dataset for Entity-Aware Medical Dialogue Generation**

Wenge Liu\*, Jianheng Tang\*, Yi Cheng\*, Wenjie Li, Yefeng Zheng, Xiaodan Liang.

*In Proceedings of International Conference on Natural Language Processing and Chinese Computing (NLPCC'2022).*

- **"My nose is running.""Are you also coughing?": Building A Medical Diagnosis Agent with Interpretable Inquiry Logics.**

Wenge Liu\*, Yi Cheng\*, Hao Wang, Jianheng Tang, Yafei Liu, Ruihui Zhao, Wenjie Li, Yefeng Zheng, Xiaodan Liang.

*In Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'2022)*

- **Guiding the Growth: Difficulty-Controllable Question Generation through Step-by-Step Rewriting**

Yi Cheng, Siyao Li, Bang Liu, Ruihui Zhao, Sujian Li, Chenghua Lin, Yefeng Zheng.

*In Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL'2021)*

- **Unifying Discourse Resources with Dependency Framework**

Yi Cheng, Sujian Li, Yueyuan Li.

*In Proceedings of International Conference on Natural Language Processing and Chinese Computing (NLPCC'2021).*

- **Zero-shot Chinese Discourse Dependency Parsing via Cross-lingual Mapping**

Yi Cheng, Sujian Li.

In *Proceedings of the 1st Workshop on Discourse Structure in Neural NLG (DSNNLG 2019)*.

- **Target-oriented Proactive Dialogue Systems with Personalization: Problem Formulation and Dataset Curation**

Jian Wang, **Yi Cheng**, Dongding Lin, Chak Tou Leong, and Wenjie Li.

In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)*.

- **RECAP: Towards Precise Radiology Report Generation via Dynamic Disease Progression Reasoning**

Wenjun Hou, **Yi Cheng\***, Kaishuai Xu\*, Wenjie Li, Jiang Liu.

*Findings of the Association for Computational Linguistics: EMNLP 2023 (EMNLP Findings 2023)*.

- **ORGan: Observation-Guided Radiology Report Generation via Tree Reasoning**

Wenjun Hou, **Yi Cheng\***, Kaishuai Xu\*, Wenjie Li, Jiang Liu.

In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2023)*.

- **Medical Dialogue Generation via Dual Flow Modeling**

Kaishuai Xu, Wenjun Hou\*, **Yi Cheng\***, Jian Wang, Wenjie Li

*Findings of the Association for Computational Linguistics: ACL 2023 (ACL Findings 2023)*.

- **CARE: Causality Reasoning for Empathetic Responses by Conditional Graph Generation.**

Jiashuo Wang, **Yi Cheng**, Wenjie Li.

*Findings of the Association for Computational Linguistics: EMNLP 2022 (EMNLP Findings 2022)*.

## RESEARCH EXPERIENCES

---

### **Towards Emotional Support Conversation Systems with Goal Awareness**

Sep 2021 – Present

Advisor: Prof. Maggie Wenjie Li, Department of Computing, The Hong Kong Polytechnic University

- Proposed a support strategy planning algorithm with lookahead heuristics to estimate the long-term effects of the adopted strategy on the user.
- Designed a dialogue framework, COOPER, which coordinates multiple specialized agents to collectively work towards a complex dialogue goal (e.g., emotional support, persuasion).
- Proposed a lightweight approach to avoid generating toxic content by manipulating the information flow within the attention layers of the language model during inference.
- Contributed to three first-author papers published at EMNLP 2022, EMNLP 2023, and AACL 2023.

### **Constructing Interpretable Medical Consultation Systems**

Feb 2021 – Dec 2021

Advisor: Dr. Yefeng Zheng, Tencent Jarvis Research Center

- Constructed the first large-scale entity-centric medical consultation dialogue dataset, MedDG.
- Proposed an interpretable medical consultation system by mimicking human doctors' inquiry logics.
- Won the 1st Prize for "the Competition at 21st China National Conference on Computational Linguistics: Automated Medical Diagnosis Systems" in 2021.
- Contributed to two first-author papers published at NLPCC 2022 and IJCAI 2022.

### **Difficulty Controllable Question Generation**

June 2020 – January 2021

Advisor: Assistant Prof. Bang Liu, Department of Computer Science and Operations Research, University of Montreal

- Defined question difficulty as the number of inference steps required to answer it and proposed a framework that progressively increases the question difficulty through step-by-step rewriting.
- Contributed to a first-author paper published at ACL 2021

### **Chinese Discourse Dependency Parsing**

Sep 2018 – June 2020

Advisor: Associate Prof. Sujian Li, Institute of Computational Linguistics, Peking University

- Proposed a zero-shot approach to conduct Chinese discourse dependency parsing via cross-lingual mapping.
- Explored the feasibility of unifying discourse resources with the dependency discourse framework.
- Contributed to two first-author papers published at DSNNLG 2019 and NLPCC 2021.

## PROFESSIONAL ACTIVITIES

---

### Challenge Organizer

March 2021 – July 2021

- Organized the ICLR 2021 Challenge: Machine Learning for Preventing and Combating Pandemics.

### Conference Reviewer

- Served as the reviewer for ACL (2022), COLING (2022), EMNLP (2021).

### Teaching Assistant

- Served as the teaching assistant for the course "COMP 6709: Advanced Natural Language Processing" in Spring 2023 at the Hong Kong Polytechnic University.

## EXTRA-CURRICULAR ACTIVITIES

---

### We Media

December 2020 – July 2022

- Served as the editor-in-chief of "夕小瑶科技说"(formerly known as "夕小瑶的卖萌屋"), an influential Chinese WeChat public account focused on the field of NLP.
- Led the team producing over 200 original articles, and the number of subscribers rose from 15k to 100k+.

## HONORS & AWARDS

---

- 1st Prize for the Competition at 21st China National Conference on Computational Linguistics: Automated Medical Diagnosis Systems 2021
- Outstanding Graduate of Peking University 2020
- "Award for Research Excellent" at Peking University 2018, 2019
- 3rd Prize for the "Tencent Cup" Program Design Competition 2018

## TECHNICAL SKILLS

---

- Language: Chinese (Native), English (Fluent)
- Standard English Test: TOEFL 111 (R: 30; L: 29; S: 27; W: 25), GRE V-156 + Q-170 + W-4.0
- Coding: Python, C/C++
- Tools: PyTorch, Tensorflow, Git, Docker, Linux, Markdown, Vim