

Chujie Zheng 郑楚杰

Email: chujiezhengchn@gmail.com

Homepage: <https://chujiezheng.github.io/>

Education

- Ph.D. candidate in Computer Science and Technology, Tsinghua University Sep 2020 – Present
Advisor: [Minlie Huang](#)
- Visiting Researcher, UCLA Nov 2023 – Jun 2024
Host: [Nanyun \(Violet\) Peng](#)
- B.S. in Foundational Mathematics and Physics, Tsinghua University Aug 2016 – Jun 2020

Work Experiences

- Researcher. Qwen Team, Alibaba Group Oct 2024 – Present
 - Built the **Qwen3** series foundation models
 - Built the **QwQ** series reasoning models
 - Built **ProcessBench** and **Qwen2.5-Math-PRM-7/72B** for process supervision in mathematical reasoning
- Research Intern. 01.AI Jul 2024 – Oct 2024
 - Built **Yi-Lightning**, which ranks **#6 on Chatbot Arena** (as of 10/14/2024)
- Research Intern. Baidu Feb 2022 – Jun 2022

Highlighted Papers

Google Scholar (**2100+** citations, h-index **23**): <https://scholar.google.com/citations?user=55zBNgUAAAAJ>

1. **ProcessBench: Identifying Process Errors in Mathematical Reasoning**
Chujie Zheng, Zhenru Zhang, Beichen Zhang, Runji Lin, Keming Lu, Bowen Yu, Dayiheng Liu, Jingren Zhou, Junyang Lin
[\[paper\]](#) [\[repo\]](#) [\[📊 data\]](#)
2. **The Lessons of Developing Process Reward Models in Mathematical Reasoning**
Zhenru Zhang, **Chujie Zheng**, Yangzhen Wu, Beichen Zhang, Runji Lin, Bowen Yu, Dayiheng Liu, Jingren Zhou, Junyang Lin
[\[paper\]](#) [\[🧠 model\]](#)
3. **Model Extrapolation Expedites Alignment**
Chujie Zheng, Ziqi Wang, Heng Ji, Minlie Huang, Nanyun Peng
[\[paper\]](#) [\[repo\]](#) [\[🧠 model\]](#) (**130K+ downloads**)
4. **On Prompt-Driven Safeguarding for Large Language Models**
Chujie Zheng, Fan Yin, Hao Zhou, Fandong Meng, Jie Zhou, Kai-Wei Chang, Minlie Huang, Nanyun Peng
ICML 2024

[\[paper\]](#) [\[repo\]](#)

5. **Large Language Models Are Not Robust Multiple Choice Selectors**

Chujie Zheng, Hao Zhou, Fandong Meng, Jie Zhou, Minlie Huang

ICLR 2024 (Spotlight: 5%)

[\[paper\]](#) [\[repo\]](#)

Other Papers

6. **Click: Controllable Text Generation with Sequence Likelihood Contrastive Learning**

Chujie Zheng, Pei Ke, Zheng Zhang, Minlie Huang

Findings of ACL 2023

7. **AugESC: Dialogue Augmentation with Large Language Models for Emotional Support Conversation**

Chujie Zheng, Sahand Sabour, Jiaxin Wen, Zheng Zhang, Minlie Huang

Findings of ACL 2023

8. **CASE: Aligning Coarse-to-Fine Cognition and Affection for Empathetic Response Generation**

Jinfeng Zhou*, Chujie Zheng*, Bo Wang, Zheng Zhang, Minlie Huang

ACL 2023 (Oral)

9. **EVA2.0: Investigating Open-domain Chinese Dialogue Systems with Large-scale Pre-training**

Yuxian Gu*, Jiaxin Wen*, Hao Sun*, Yi Song, Pei Ke, Chujie Zheng, Zheng Zhang, Jianzhu Yao, Lei Liu, Xiaoyan Zhu, Minlie Huang

Machine Intelligence Research 2023

10. **CDConv: A Benchmark for Contradiction Detection in Chinese Conversations**

Chujie Zheng*, Jinfeng Zhou*, Yinhe Zheng, Libiao Peng, Zhen Guo, Wenquan Wu, Zhengyu Niu, Hua Wu, Minlie Huang

EMNLP 2022 (Oral)

11. **COLD: A Benchmark for Chinese Offensive Language Detection**

Jiawen Deng*, Jingyan Zhou*, Hao Sun, Chujie Zheng, Fei Mi, Helen Meng, Minlie Huang

EMNLP 2022 (Oral)

12. **On the Safety of Conversational Models: Taxonomy, Dataset, and Benchmark**

Hao Sun*, Guangxuan Xu*, Jiawen Deng, Jiale Cheng, Chujie Zheng, Hao Zhou, Nanyun Peng, Xiaoyan Zhu, Minlie Huang

Findings of ACL 2022

13. **CEM: Commonsense-aware Empathetic Response Generation**

Sahand Sabour, Chujie Zheng, Minlie Huang

AAAI 2022 (Oral)

14. **Exploring Prompt-based Few-shot Learning for Grounded Dialog Generation**

Chujie Zheng, Minlie Huang

arXiv:2109.06513

15. **EVA: An Open-Domain Chinese Dialogue System with Large-Scale Generative Pre-Training**
Hao Zhou*, Pei Ke*, Zheng Zhang*, Yuxian Gu, Yinhe Zheng, **Chujie Zheng**, Yida Wang, Chen Henry Wu,
Hao Sun, Xiaocong Yang, Bosi Wen, Xiaoyan Zhu, Minlie Huang, Jie Tang
arXiv:2108.01547
16. **Towards Emotional Support Dialog Systems**
Siyang Liu*, **Chujie Zheng***, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, Minlie Huang (*:
Equal contribution)
ACL-IJCNLP 2021 (Oral)
17. **CoMAE: A Multi-factor Hierarchical Framework for Empathetic Response Generation**
Chujie Zheng, Yong Liu, Wei Chen, Yongcai Leng, Minlie Huang
Findings of ACL-IJCNLP 2021
18. **PsyQA: A Chinese Dataset for Generating Long Counseling Text for Mental Health Support**
Hao Sun*, Zhenru Lin*, **Chujie Zheng**, Siyang Liu, Minlie Huang
Findings of ACL-IJCNLP 2021
19. **Difference-aware Knowledge Selection for Knowledge-grounded Conversation Generation**
Chujie Zheng, Yunbo Cao, Daxin Jiang, Minlie Huang
Findings of EMNLP 2020
20. **KdConv: A Chinese Multi-domain Dialogue Dataset Towards Multi-turn Knowledge-driven Conversation**
Hao Zhou*, **Chujie Zheng***, Kaili Huang, Minlie Huang, Xiaoyan Zhu
ACL 2020
21. **ChID: A Large-scale Chinese IDiom Dataset for Cloze Test**
Chujie Zheng, Minlie Huang, Aixin Sun
ACL 2019

Selected Awards and Honors

- | | |
|--|-------------|
| • Comprehensive Merit Scholarship, Tsinghua University | 2021 – 2024 |
| • Outstanding Undergraduate, Tsinghua University | 2020 |
| • China National Scholarship (Top 2/100) | 2019 |
| • Comprehensive Merit Scholarship, Tsinghua University | 2018 |

Academic Services

- **Area Chair:** ACL (24/25), EMNLP (24/25), NAACL (25), ACL Rolling Review (24/25)
- **Reviewer:** ICLR (25), NeurIPS (24/25), ICML (24), COLM (24/25), ACL (22/23), EMNLP (21/22), NAACL (24), EACL (23), ACL Rolling Review (21/22/23), CogSci (24), AAAI (22/23)