

# Haozhe Ji (计昊哲)

Email: [jihaozhe@gmail.com](mailto:jihaozhe@gmail.com)

Web: [haozheji.github.io](https://haozheji.github.io)

Phone: +86 15652655811

## RESEARCH INTERESTS

My research interests center around developing **theoretically grounded and scalable algorithms** to improve neural language models on **natural language generation** and **language model alignment**. Specifically, my current research focused to address the distribution misspecification due to the limited expressivity of Auto-Regressive Modeling (ARM) and the inherent bias of Maximum Likelihood Estimation (MLE).

- To overcome the expressivity limitation of ARM, my research considered a broader spectrum of expressive model families, including semi-parametric models [8,9], memory-augmented models [5], latent variable models [6] and energy-based models [2].
- To tackle the inherent bias of MLE, my research proposed theoretically grounded and practically accessible training objectives [3, 1] and decoding frameworks [2], aiming to achieve better alignment with human language.

## EDUCATION

**Tsinghua University**, Beijing, China  
*Ph.D. Student*, Computer Science and Technology  
*Advisor*: Minlie Huang

September 2020 - Present

**Tsinghua University**, Beijing, China  
*B.E.*, Electronic Engineering

September 2016 - July 2020

## PREPRINTS

## PUBLICATIONS

- [1] **Towards Efficient and Exact Optimization of Language Model Alignment**  
**Haozhe Ji**, Cheng Lu, Yilin Niu, Pei Ke, Hongning Wang, Jun Zhu, Jie Tang, Minlie Huang  
*International Conference on Machine Learning (ICML)*, 2024.
- [2] **Language Model Decoding as Direct Metrics Optimization**  
**Haozhe Ji**, Pei Ke, Hongning Wang, Minlie Huang  
*International Conference on Learning Representations (ICLR)*, 2024.
- [3] **Tailoring Language Generation Models under Total Variation Distance**  
**Haozhe Ji**, Pei Ke, Zhipeng Hu, Rongsheng Zhang, Minlie Huang  
*International Conference on Learning Representations (ICLR)*, 2023. (**Notable top 5%**)
- [4] **Curriculum-Based Self-Training Makes Better Few-Shot Learners for Data-to-Text Generation**  
Pei Ke, **Haozhe Ji**, Zhenyu Yang, Yi Huang, Junlan Feng, Xiaoyan Zhu, Minlie Huang  
*International Joint Conference on Artificial Intelligence (IJCAI)*, 2022.
- [5] **LaMemo: Language modeling with look-ahead memory**  
**Haozhe Ji**, Rongsheng Zhang, Zhenyu Yang, Zhipeng Hu, Minlie Huang  
*North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022. (**Oral**)
- [6] **DiscoDVT: Generating Long Text with Discourse-Aware Discrete Variational Transformer**  
**Haozhe Ji**, Minlie Huang  
*Empirical Methods in Natural Language Processing (EMNLP)*, 2021. (**Oral**)

[7] **Jointgt: Graph-text joint representation learning for text generation from knowledge graphs**  
 Pei Ke, **Haozhe Ji**, Yu Ran, Xin Cui, Liwei Wang, Linfeng Song, Xiaoyan Zhu, Minlie Huang  
*Findings of the Association for Computational Linguistics (Findings of ACL)*, 2021.

[8] **Language generation with multi-hop reasoning on commonsense knowledge graph**  
**Haozhe Ji**, Pei Ke, Shaohan Huang, Furu Wei, Xiaoyan Zhu, Minlie Huang  
*Empirical Methods in Natural Language Processing (EMNLP)*, 2020. (*Oral*)

[9] **Generating commonsense explanation by extracting bridge concepts from reasoning paths**  
**Haozhe Ji**, Pei Ke, Shaohan Huang, Furu Wei, Minlie Huang  
*Asia-Pacific Chapter of the Association for Computational Linguistics (AACL)*, 2020.

[10] **Sentilare: Linguistic knowledge enhanced language representation for sentiment analysis**  
 Pei Ke\*, **Haozhe Ji**\*, Siyang Liu, Xiaoyan Zhu, Minlie Huang  
*Empirical Methods in Natural Language Processing (EMNLP)*, 2020.

[11] **Denoising distantly supervised open-domain question answering**  
 Yankai Lin, **Haozhe Ji**, Zhiyuan Liu, Maosong Sun  
*Annual Meeting of the Association for Computational Linguistics (ACL)*, 2018.

RESEARCH EXPERIENCE	<div> <div> <b>CoAI Lab, Tsinghua University</b>  <i>Ph.D. Candidate (Supervisor: <a href="#">Minlie Huang</a>)</i> </div> <div> September 2020 - July 2025 (Expected) </div> </div> <div> <div> <b>Natural Language Comupting group, Microsoft Research Asia</b>  <i>Research Intern (Supervisors: <a href="#">Shaohan Huang</a>, <a href="#">Furu Wei</a>)</i> </div> <div> July 2019 - July 2020 </div> </div>
SERVICES	<div> <b>Reviewer/Program Committee:</b> ACL, EMNLP, NAACL, ARR </div>
AWARDS	<div> <div> <b>Tang Junyuan (唐君远) Scholarship</b>, Tsinghua University </div> <div> 2023 </div> </div> <div> <div> <b>Sohu Scholarship</b>, Tsinghua University </div> <div> 2022 </div> </div> <div> <div> <b>Yang Huiyan (杨惠妍) Scholarship</b>, Tsinghua University </div> <div> 2021 </div> </div> <div> <div> <b>Comprehensive Merit Scholarship</b>, Tsinghua University </div> <div> 2019 </div> </div> <div> <div> <b>Comprehensive Merit Scholarship</b>, Tsinghua University </div> <div> 2017 </div> </div> <div> <div> <b>Gold Medal</b>, 32nd China Physics Olympiads (CPhO) </div> <div> 2015 </div> </div> <div> <div> <b>Distinguished Honor Roll</b> (Top 1%), American Mathematics Contest 12 (AMC12) </div> <div> 2015 </div> </div>