YU ZHANG

Soochow University

Institute of Artificial Intelligence

School of Computer Science & Technology

Suzhou, China

📞 (+86) 15050769956

✉ yzhang.cs@outlook.com

⚙ https://github.com/yzhangcs

🏠 https://yzhang.site

PROFILE

I am currently a third-year PhD student at HLT@SUDA, advised by Prof. Guohong Fu. I expect to graduate in 2025. Prior to this, I received my M. Eng. degree from Soochow University in 2021.

My early research focused on structured prediction tasks, specifically dependency parsing and constituency parsing. Currently, my research interests have evolved to focus on developing efficient text generation models. I am particularly intrigued by the prospect of developing hardware-efficient methods for linear-time sequence modeling. As a disciple of parallel programming, I am passionate about exploring techniques that harness the power of parallel computing to develop scalable subquadratic models.

PUBLICATIONS

[Semantic Scholar] [Google Scholar] [DBLP] (* denotes equal contributions)

- Gated Slot Attention for Efficient Linear-Time Sequence Modeling
  Yu Zhang*, Songlin Yang*, Rui-Jie Zhu, Yue Zhang, Leyang Cui, Yiqiao Wang, Bolun Wang, Freda Shi, Bailin Wang, Wei Bi, Peng Zhou, Guohong Fu
  NeurIPS 2024

- Parallelizing Linear Transformers with the Delta Rule over Sequence Length
  Songlin Yang, Bailin Wang, **Yu Zhang**, Yikang Shen, Yoon Kim
  NeurIPS 2024

- Scalable MatMul-free Language Modeling
  Rui-Jie Zhu, **Yu Zhang**, Ethan Sifferman, Tyler Sheaves, Yiqiao Wang, Dustin Richmond, Peng Zhou, Jason K. Eshraghian

- Non-autoregressive Text Editing with Copy-aware Latent Alignments
  Yu Zhang*, Yue Zhang*, Leyang Cui, Guohong Fu
  EMNLP 2023

- Semantic Role Labeling as Dependency Parsing: Exploring Latent Tree Structures Inside Arguments
  Yu Zhang, Qingrong Xia, Shilin Zhou, Yong Jiang, Guohong Fu, Min Zhang
  COLING 2022

- Fast and Accurate End-to-End Span-based Semantic Role Labeling as Word-based Graph Parsing

Shilin Zhou, Qingrong Xia, Zhenghua Li, **Yu Zhang**, Yu Hong, Min Zhang
COLING 2022 (*Best Paper Award*)

- Fast and Accurate Neural CRF Constituency Parsing
  Yu Zhang*, Houquan Zhou*, Zhenghua Li
  IJCAI 2020

- Efficient Second-Order TreeCRF for Neural Dependency Parsing
  Yu Zhang, Zhenghua Li, Min Zhang
  ACL 2020

- Is POS Tagging Necessary or Even Helpful for Neural Dependency Parsing?
  Houquan Zhou*, **Yu Zhang***, Zhenghua Li, Min Zhang
  NLPCC 2020 (*Best Paper Award*)

- HLT@SUDA at SemEval-2019 Task 1: UCCA Graph Parsing as Constituent Tree
  Parsing
  Wei Jiang, Zhenghua Li, **Yu Zhang**, Min Zhang
  SemEval 2019

PROJECTS

- **FLA**: A Triton-Based Library for Hardware-Efficient Implementations of Linear
  Attention Mechanism.

- **SuPar**: State-of-the-art syntactic/semantic parsers (760 stars, 21k downloads)
  A Python package designed for structured prediction, including reproductions of
  many state-of-the-art syntactic/semantic parsers (with pretrained models for more
  than 19 languages), and highly-parallelized implementations of several well-known
  structured prediction algorithms.

EDUCATION

2018.9-2021.6:   *Master of Computer Science* from **Soochow University**
2014.9-2018.6:   *Bachelor of Engineering* from **Soochow University**

EXPERIENCE

- 2021 - present. PhD student at Soochow University, advised by Prof. Guohong Fu.

- 2023 - 2024. Research Intern at Tencent AI Lab, mentored by Wei Bi.

- 2020 - 2021. Research Intern at Alibaba DAMO Academy, mentored by Yong Jiang.

- 2018 - 2021: M.S. student at Soochow University, advised by Prof. Zhenghua Li.