

YAO FU

+44 7564 053184

<https://franxyao.github.io>

<https://github.com/FranxYao>

yao.fu@ed.ac.uk

RESEARCH AREA

Large Language Models

Generative Models

Latent Variable Models

Complex Reasoning

Emergent Abilities

Controllable Generation

ACADEMIC EXPERIENCES

University of Edinburgh, School of Informatics, Edinburgh, UK

Ph.D. in Computer Science. Advisor: [Prof. Mirella Lapata](#)

Sep 2020 - Dec 2023

Cornell University, Tech Campus, New York, NY

Visiting Research Student. Advisor: [Prof. Alexander Rush](#)

Sep 2019 - Sep 2020

Columbia University, Department of Computer Science, New York, NY

M.S. in Computer Science. Advisor: [Prof. John Cunningham](#)

Aug 2018 - Dec 2019

Peking University, Department of Computer Science, Beijing, CN

B.S. in Computer Science. Advisor: [Prof. Yansong Feng](#)

Sep 2013 - Jun 2018

INDUSTRIAL EXPERIENCES

Google DeepMind. Mountain View.

(Incoming) Sep - Dec 2023

Research Intern in Training Large Language Models. With Dr. [Mahesh Sathiamoorthy](#)

MIT-IBM Watson AI Lab. Cambridge.

Jun - Sep 2023

Research Intern in Training Large Language Models. With Dr. [Rameswar Panda](#) and Prof. [Yoon Kim](#)

Allen Institute for AI. Seattle.

Jul - Dec 2022

Research Intern in Large Language Models Reasoning. With Dr. [Tushar Khot](#), Dr. [Ashish Sabharwal](#) and Prof. [Hao Peng](#)

Alibaba DAMO Academy. Hangzhou.

Jan - Jul 2020

Research Intern in Structured Prediction and Text Generation. With Dr. [Chuanqi Tan](#) and Dr. [Mosha Chen](#)

Tencent AI Lab. Seattle.

May - Aug 2019

Research Intern in Structured Prediction. With Dr. [Kun Xu](#) and Dr. [Dong Yu](#)

Bytedance AI Lab. Beijing.

Jan - Aug 2018

Research Intern in Sequence Modeling and Sentence Generation. With Prof. [Lei Li](#) and Prof. [Hao Zhou](#)

SELECTED WORK

17. Improving Language Model Negotiation with Self-Play and In-Context Learning from AI Feedback. 2023

Yao Fu, Hao Peng, Tushar Khot, and Mirella Lapata

Two language models negotiate with each other and continuously improve their negotiation strategies by multi-round game playing and iterative in-context learning from AI feedback.

16. Specializing Smaller Language Models towards Multi-Step Reasoning. ICML 2023

Yao Fu, Hao Peng, Litu Ou, Ashish Sabharwal, and Tushar Khot

Improving small language models' math chain-of-thought ability by trading their generic abilities.

15. Complexity-Based Prompting for Multi-Step Reasoning. ICLR 2023

Yao Fu, Hao Peng, Ashish Sabharwal, Peter Clark and Tushar Khot

State-of-the-art reasoning performance on math word problems by prompting GPT with complex reasoning chains.

14. How does GPT Obtain its Ability? Tracing Emergent Abilities of Language Models to their Sources. 2022

Yao Fu

[Hacker News Top 3 Trending](#). Analysing sources of emergent abilities of Large Language Models from first principle. ([link](#))

CONFERENCE PUBLICATIONS AND PREPRINTS

13. C-Eval: A Multi-Level Multi-Discipline Chinese Evaluation Suite for Foundation Models. Arxiv 2023

Yuzhen Huang, Yuzhuo Bai, Zhihao Zhu, Junlei Zhang, Jinghan Zhang, Tangjun Su, Junteng Liu, Chuancheng Lv, Yikai Zhang, Jiayi Lei, Yao Fu, Maosong Sun, Junxian He

A evaluation suite consisting of 52 subjects of STEM/ Social science/ Humanity/ Other testing language models' Chinese ability (knowledge and reasoning).

12. To Repeat or Not To Repeat: Insights from Scaling LLM under Token-Crisis. Arxiv 2023

Fuzhao Xue, Yao Fu, Wangchunshu Zhou, Zangwei Zheng, Yang You

Investigating the effect of repeating tokens during pretraining large language models.

11. Decomposed Prompting: A Modular Approach for Solving Complex Tasks. ICLR 2023

Tushar Khot, Harsh Trivedi, Matthew Finlayson, Yao Fu, Kyle Richardson, Peter Clark and Ashish Sabharwal

Decomposing complex task into simpler sub-tasks then solve each of them by prompting language models.

10. Scaling Structured Inference with Randomization. ICML 2022

Yao Fu, John Cunningham and Mirella Lapata

A family of randomized dynamic programming algorithms for scaling up the inference of graphical models.

9. Data-to-Text Generation with Variational Sequential Planning. TACL 2022.

Ratish Puduppully, Yao Fu and Mirella Lapata

A sequential planning model for generating long documents given structured input data.

8. Noisy-Labeled NER with Confidence Estimation. NAACL 2021.

Kun Liu*, Yao Fu*, Chuanqi Tan, Mosha Chen, Ningyu Zhang, Songfang Huang and Sheng Gao (*equal contribution)

A method for estimating the confidence of annotations under a noisy labeled learning setting.

7. Probing BERT in Hyperbolic Spaces. ICLR 2021.

Boli Chen*, Yao Fu*, Guangwei Xu, Pengjun Xie, Chuanqi Tan, Mosha Chen and Liping Jing (*equal contribution)

A probing model for finding hierarchical information from contextualized representations in Hyperbolic spaces.

6. Prototypical Representation Learning for Relation Extraction. ICLR 2021.

Ning Ding, Xiaobin Wang, Yao Fu, Guangwei Xu, Rui Wang, Pengjun Xie, Ying Shen, Fei Huang, Hai-Tao Zheng and Rui Zhang

A contrastive learning method for pretraining relation prototype representations on hyperspheres.

5. Nested Named Entity Recognition with Partially Observed TreeCRFs. AAAI 2021.

Yao Fu*, Chuanqi Tan*, Mosha Chen, Songfang Huang and Fei Huang. (*Equal contribution)

A new masked Inside algorithm for learning partially observed TreeCRF.

4. Latent Template Induction with Gumbel-CRF. NeurIPS 2020.

Yao Fu, Chuanqi Tan, Mosha Chen, Bin Bi and Alexander M. Rush

A Gumbel-softmax based continuous relaxation method for reparameterizing structured VAEs for controllable sentence generation.

3. Paraphrase Generation with Latent Bag of Words. NeurIPS 2019.

Yao Fu, Yansong Feng and John Cunningham

An interpretable model with differentiable content planning and surface realization by gumbel-topk reparameterization.

2. Rethinking Text Attribute Transfer: A Lexical Analysis. INLG 2019.

Yao Fu, Hao Zhou, Jiaze Chen and Lei Li

An empirical study of existing text style transfer datasets and models to reveal critical changes during text style transfer.

1. Natural Answer Generation with Heterogeneous Memory. NAACL 2018.

Yao Fu and Yansong Feng

A key-value memory network for retrieving knowledge from various sources then generating answer sentences.

WORKSHOP PUBLICATIONS

1. Just DREAM about it: Figurative Language Understanding with DREAM-FLUTE. The Third Workshop on Figurative Language Processing. In conjunction with EMNLP 2022

Yuling Gu, **Yao Fu**, Valentina Pyatkin, Ian Magnusson, Bhavana Dalvi Mishra and Peter Clark

Ranked top 1 in the task leaderboard. A mental model utilizing scene elaboration for understanding figurative language.

BLOG AND OPEN SOURCE PROJECTS

7. Towards Complex Reasoning: the Polaris of Large Language Models

A roadmap towards building language models of strong reasoning capabilities. Covers the full development stages: pretraining, continue training, supervised finetuning, reinforcement learning, chain-of-thought prompting, and evaluation. ([link](#))

6. ChatArena: Multi-Agent Language Game Environments for Large Language Models (700+ stars)

A library that provides multi-agent language game environments and facilitates research about autonomous LLM agents and their social interactions. ([link](#))

5. Chain of Thought Hub: Measuring LLMs' Reasoning Performance (600+ stars)

Benchmarking Large Language Models reasoning abilities. ([link](#))

4. A Closer Look at Large Language Models Emergent Abilities

Scrutinizing Large Language Model Emergent Abilities to see if the paradigm is really shifting to in-context learning. ([link](#))

3. Deep Generative Models for Natural Language Processing (300+ stars)

A roadmap tracking past, present, and future about generative models for NLP. ([link](#))

2. Why S4 is Good at Long Sequence: Remembering a Sequence with Online Function Approximation

A blog post explaining the S4 model with function approximation theory. ([link](#))

1. How to Write Variational Inference and Generative Models for Natural Language Processing

A tutorial about variational inference and deep generative models for NLP. ([link](#))

INVITED TALKS

10. Pretraining, Instruction Tuning, Alignment: Towards Building Large Language Models from First Principles

University of Cambridge

May 2023

Microsoft Research. Cambridge

May 2023

MIT-IBM Watson AI Lab. Boston

Mar 2023

University of Edinburgh. Edinburgh

Mar 2023

University College London. London

Mar 2023

New York University.

Feb 2023

Tsinghua University. Beijing

Dec 2022

Fudan University. Shanghai

Dec 2022

9. A Closer Look at Emergent Abilities of Large Language Models

Bytedance. Beijing

Dec 2022

8. Challenges and Opportunities in Large Language Model Regime

MIT-IBM Watson AI Lab. Cambridge

Oct 2022

7. Challenges in Encoding Long Documents.

Microsoft Research. Redmond

Aug 2022

Microsoft Research. Beijing

Sep 2022

6. Scaling Structured Inference with Randomization.

Tsinghua University. Beijing

Aug 2022

AI TIMES. Online

Jul 2022

Language at Edinburgh. Edinburgh

Mar 2022

5. Deep Structured Prediction: Inference, Reparameterization and Applications

Bytedance AI Lab. Beijing

Jun 2021

4. Language Technologies, Artificial Intelligence and their Implications

Huawei. Berlin

May 2021

3. Controllable Text Generation with Structured Latent Variables

Alibaba. Hangzhou

Sep 2020

2. Learning Structured Representations for Text

Peking University. Beijing

Sep 2020

Westlake Institute of Advanced Study. Hangzhou

Sep 2020

1. On the Evaluation of Text Generation Models

Interactions Inc. New York

Sep 2019

TEACHING

5. Natural Language Understanding, Generation, and Machine Translation

University of Edinburgh. Teaching Assistant. Edinburgh

Spring 2022, 2023

4. Probabilistic Modeling and Reasoning

University of Edinburgh. Teaching Assistant. Edinburgh

Fall 2021

3. Text Generation with Neural Sequence Models

Peking University. Guest Lecture. Beijing

Spring 2021, 2022

2. Advanced Probabilistic Machine Learning Seminar

Alibaba DAMO Academy. Instructor. Hangzhou

Spring 2020

1. Applied Machine Learning

Columbia University. Course Assistant. New York

Spring 2019

SERVICE

ACL 2023 Student Research Workshop

Student Chair

2022 - 2023

Edinburgh ILCC Poster Session

Organizer

Jun 2022

Edinburgh ILCC Seminar. Statistical Learning Theory

Organizer

Summer 2021

Edinburgh ILCC Seminar. *Combinatorial Categorical Grammar*

Organizer

Spring 2021

ICML 2022

Reviewer

NeurIPS 2021-2023

Reviewer

ICLR 2021-2023

Reviewer

ACL 2019-2022

Reviewer

EMNLP 2018-2022

Reviewer

NAACL 2019-2021

Reviewer

EACL 2021

Reviewer

IJCAI 2021-2022

Reviewer

NeurIPS 2022 Math AI Workshop

Reviewer

ACL 2023 Natural Language Reasoning and Structured Explanations workshop

Reviewer

HONORS

Alibaba Outstanding Intern of the Year 2020

Alibaba Group

Dec 2020

Outstanding Undergraduate Dissertation Award

Peking University

Jun 2018

Elite Collegiate Award

China Computer Federation

Jul 2017

C.X. Zhong & X.N. Xia Scholarship

Peking University

Jun 2015, 2016, 2017

Scholarship of Exchange Program

China Scholarship Council

Feb 2016