Shiyu Zhao

✉ sy-zhao19@mails.tsinghua.edu.cn◇ ☎ +86 15009500270 ◇ https://shiyu-zhao.netlify.app/

EDUCATION

Tsinghua University, China *Sep 2019 - Present*

Computer Science and Technology

- Department: Institute for Interdisciplinary Information Sciences (IIIS)
- Class: Yao Class, a pilot class led by Turing award winner, Andrew Chi-Chih Yao
- GPA: 3.92/4.0
- Admitted as the highest scorer among 70,000+ students in Chinese College Entrance Examination | Ningxia
- Relevant coursework: Machine Learning(4.0), Artificial Intelligence: Principles and Techniques(4.0), Introduction to Data Science(4.0), Mathematics for Computer Science(4.0), Fundamentals of Programming(4.0), Research Immersion Training(4.0)...

RESEARCH INTERESTS

Graph Representation Learning; Natural Language Processing; Computational Biology; Computation Social Science; Cognition and Reasoning;

PUBLICATIONS

Xiao Liu*, **Shiyu Zhao***, Kai Su*, Yukuo Cen, Jiezhong Qiu, Mengdi Zhang, Wei Wu, Yuxiao Dong, Jie Tang et al. "Mask and Reason: Pre-Training Knowledge Graph Transformers for Complex Logical Queries". In: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.* KDD '22. [link] (∗ **indicates equal contribution)**

RESEARCH EXPERIENCE

LogicGNN: Logic Message Passing Graph Neural Network

Mentor: Prof. Jian Tang, Prof. Leon Bergen *May 2022 - present*

Montreal Institute for Learning Alogorithms (MILA), University of California San Diego *Planned for ICML 2023*

- Proposes the first **systematically generalizable** and **scalable** reasoning model on KG.
- Proposes to model **one-step logic inference** as **triangle update** on graph inspired by logic programming and formalize it under the designed GNN framework.
- Identify **reasonable patterns** and conduct partial reasoning with the help of **auxiliary edges** on graphs.
- Already achieved SOTA in link prediction, planned for ICML 2023.

End-to-end Small Molecule Entity Discovery

Mentor: Prof. Jian Tang *March 2022 - May 2022*

Montreal Institute for Learning Alogorithms (MILA) *Package under development*

- Proposes to use a encoder and a retriever to retrieve **similar** molecules out of the whole dataset as **supportive evidence**
- Introduces a **refinement stage** to utilize only supportive evidence for downstream task.
- Builds a **model-agnostic enhancement** pipeline for small molecule property prediction.
- Achieved positive results on HIV, ClinTox and PCQM, will be sealed as package.

Mask and Reason: Pre-Training Knowledge Graph Transformers for Complex Logical Queries

Xiao Liu, **Shiyu Zhao***, *Kai Su*, Yukuo Cen, Jiezhong Qiu, Mengdi Zhang, Wei Wu, Yuxiao Dong, Jie Tang*

Mentor: Prof. Jie Tang, Prof. Yuxiao Dong *Jan 2021 - Nov 2021*

Knowledge Engineering Group (KEG), CS Department, Tsinghua University *Accepted to KDD 2022*

- Introduces a well-functioning **pretrain-finetune paradigm** into knowledge graph area with generalizability and interpretability.
- Creatively design a **KG triple transformation** method to enable transformer to handle KG elegantly and a mechanism to **unify different tasks** of knowledge graph problems.

- Achieves **SOTA** on both in-domain and out-of-domain reasoning task, **significantly outperforms** previous SOTA CQD(ICLR 2021 best paper) by over $\mathbf{12.1}\%$ relatively on FB15k-237 and over $\mathbf{6.4}\%$ relatively on NELL995.
- Accepted by KDD 2022.

Graph Reasoning with GNN Matching
Mentor: Prof. Jie Tang *Oct 2020 - present*
Knowledge Engineering Group (KEG), CS Department, Tsinghua University *ongoing*

- Proposes a model based on **graph matching** to extract **universal human-readable rules** of relations over large knowledge graphs; provides some insights of **explainability**.
- Builds up some **theory guarantee** of the capability of the method
- Cooperates with others in building **a large knowledge graph system** over Wikidata. (ongoing)

PROJECTS

GLUE+: Comprehensive and distinguishable graph-linked embedding for multi-omics single-cell data integration

- Solve the **indistinguishability** of aggregating multi-omics data on the graph for the graph-linked embedding
- Enrich the multi-omics information of graph embedding by using **multiple aggregators** in the GNN
- Gained enhancement over the original **GLUE** model

Improvement of Random Matrix Factorization of Large-scale Network Embedding

- Improves NetMF embedding algorithm by **single-view SVD**, avoids storage of dense matrix and saves space.
- Sets the **decay rate** of singular value and uses **freigs** algorithm to speed up the factorization, achieves linear bound.

A Survey on Non-black-box Simulator of Zero-knowledge Interactive proofs

- Surveys over **FLS-type protocal** and its application over non-black-box simulator in zero-knowledge proof.
- Gains insights of the power of non-black-box proofs over some black-box problems.

SKILLS

- Programming skills: Python, PyTorch(proficient), SQL, C /C++, bash, MATLAB, LaTeX, Verilog, Go
- Language skills: Chinese(native), English(GRE: 338: V169, Q169, W4.0, TOEFL 108: R30 L26 S25 W27)

HONORS AND AWARDS

- First place in Chinese College Entrance Examination | Ningxia *2019*
  - rank: 1/71702, grades: 697/725
- The first prize scholarship for Freshmen | Tsinghua University *2019*
- Dean's list | Tsinghua University *2020*
- Dean's list | Tsinghua University *2021*
- Dean's list | Tsinghua University *2022*

LEADERSHIP

- Class Leader | Yao Class 92, Tsinghua University *2019-2020*
- Member of Organization Department of Student Union | IIIS, Tsinghua University *2020-2021*
- Member of Science and Technology Department of Student Union | IIIS, Tsinghua University *2020-2021*
- Class Leader | Yao Class 92, Tsinghua University *2021-2022*
- Minister of Literature and Art Department of the Student Union | IIIS, Tsinghua University *2021-2022*
- Class Leader | Yao Class 92, Tsinghua University *2022-present*