Kun Gao

Contact Information

Email: kungao@pku.edu.cn Mobile: (+65) 8694-4563 Web: Personal Page

Education

Peking University (PKU), Ph.D. in Computer Software and Theory, Beijing, China July 2018 – July 2023

Advisors: Prof. Hanpin Wang, Prof. Yongzhi Cao

Courses: Machine Learning, Knowledge Representation, Software Analyze, Formal Language and Automata, etc.

Research Topic: Neuro-Symbolic Inductive Logic Programming

Dissertation: Research on Inductive Logic Programming with Neuro-Symbolic Methods

National Institute of Informatics (NII), Researcher Intern, Tokyo, Japan Sept. 2019 – Mar. 2020

Advisors: Prof. Katsumi Inoue

Research Topic: Knowledge Representation and Reasoning, Inductive Logic Programming

University of Science and Technology Beijing (USTB), B.E. in Computer Science, Beijing, China Sept. 2014 – June 2018

GPA: 3.75/4; Rank: 2/68

Courses: Computer Organization, Operating System, The Mathematic Foundation for Information Security, Assemble Language, Data Structure and Algorithm, etc.

Research Experience

Differentiable Inductive Logic Programming (NII, PKU) Aug. 2019 – Now

- **Knowledge retrieval from knowledge graphs and time series**. Learning first-order rules from knowledge graphs needs a model that has scalability. In addition, lots of first-order rule learning models need strong language bias as input, which is a type of domain knowledge and increases the useability gap of the models. To design a scalable model without strong language bias as input, I design a sampling algorithm, a propositionalization method, and an interpretable neural network framework to learn first-order rules from knowledge graphs. I tested the learned rules on link prediction tasks. The research work has been published in IJCAI and Artificial Intelligence journal.

- **Knowledge representation**. Learning proportional rules from noisy data is still hard to obtain a good performance. Besides, utilizing existing background knowledge represented by rules for predictions is challenging for the existing rule-learning method. To solve the above challenges, I design a matrix representation method of logic programs and an algebraic logic programming method. Further, I design an interpretable neural network to learn rules from data. In addition, the interpretable neural network can use existing symbolic rules to learn more rules. The work has been published in Machine Learning journal.

- **Mathematics foundation**. In the research, I have learned the foundation of logic programming, the symbolic and non-symbolic inductive logic programming methods, logic programming software such as Prolog, embedding-based relational learning models, and probabilistic logic programming models such as Markov Logic Networks.

- **Induction reasoning with LLMs**. To utilize the learning ability of large language models (LLMs), I design a method to perform inductive tasks with LLMs. The experiment results indicate LLMs can learn simple relations based on a limited number of facts as input.

Blockchain and Cloud Storage (PKU) Aug. 2018 – Aug.2019

- **Formal verification**. To use formal language to verify the correctness of smart contracts in blockchain systems and cloud storage systems, I have learned the data structure and concepts in blockchain systems, the architect of block-based cloud storage systems, formal languages, and separation logic.

Working Experience

Institute of High Performance Computing, Agency for Science, Technology, and Research, Singapore

Research Scientist Aug. 2023 – Now

Projects and Research Topics:

- **Transparency AI**. To learn the semantics of multimodality data such as time series data and image data, I have performed research on learning first-order rules from time series and image data. I use the feature learning method and define possible predicates that describe the temporal, spatial, and attribute characters of a single or several features. Then, I design a differentiable inductive logic programming model to generalize first-order rules from the above proceed data. Some preliminary works have been submitted to serval conferences. I am still on the way to inducing more interpretable rules from data and preparing for top-tier conferences and journals.
- **Deduction reasoning for LLMs**. To improve the reasoning ability of LLM, I design a method to transfer the prior knowledge into the LLM model to improve the deduction ability of LLM.
- **Explainable AI**. To understand the predictions of neural networks such as LSTM, RNN and transformer, I adopt post-hoc explainable AI methods such as **SHAP** and LIME to obtain global and local explanation results among the existing features in tabular data and time series data.

Data, Knowledge, and Intelligence team, Microsoft Research Asia (MSRA), Beijing, China

Research Scientist Intern June 2022 – Oct. 2022

Research Topic: Using **self-attention** and **transformer** to learn interpretable rules from industry time series data in cloud source allocation problems.

Intel, Beijing, China

Software Engineer Intern Sept. 2017 – Mar. 2018

Job Scope: Maintain **cloud storage** system such as OpenStack Swift with Python.

Publications

1. **K. Gao**, K. Inoue, Y. Cao, and H. Wang (2024): A differentiable first-order rule learner for inductive logic programming. *Artificial Intelligence*, 331, 104108. (IF: 14.4)
2. **K. Gao**, K. Inoue, Y. Cao, and H. Wang (2022): Learning First-Order Rules with Differentiable Logic Program Semantics. IJCAI 2022: 3008-3014. (Long Presentation, rank **top 3.75%** among IJCAI 2022 accepted papers.)
3. **K. Gao**, H. Wang, Y. Cao, and K. Inoue (2022). Learning from interpretation transition using differentiable logic programming semantics. *Machine Learning*, 111(1), 123-145. (IF: 7.5)

4. C. Yang, L. Wang, **K. Gao**, and S. Li (2023). Reinforcement Logic Rule Learning for Temporal Point Processes. arXiv preprint arXiv:2308.06094.
5. Z. Jin, H. Wang, L. Zhang, B. Zhang, **K. Gao**, and Y. Cao (2019). Reasoning about Block-based Cloud Storage Systems. arXiv preprint arXiv:1904.04442.

Skills

Programming languages: Python, C++, SQL, Java, PHP, Assembly Language, Prolog, etc.

Softwares: TensorFlow, PyTorch, MATLAB, LaTeX, Linux, Wireshark, Database, Adobe Illustrator, etc.

Research topics: Neuro-Symbolic Methods, Knowledge Representation and Learning, Explainable AI, Reasoning ability for LLMs.

Awards

IJCAI-24 Conference Grant	2024
National Award at PKU	2022
Excellent Merit Student at PKU and USTB	2022, 2017, 2016
Outstanding Graduates at USTB	2018
Top-Class Scholarships at USTB	2016
Programming Contest Third-Class Award at USTB	2015

Academic Services

Peer-reviewed Journals

PC Member, ECAI-24	2024
ACM Transactions on Intelligent Systems and Technology	2024
IEEE Transactions on Knowledge and Data Engineering	2024

Presentations

IJCAI-24 Long presentation

K. Gao, K. Inoue, Y. Cao, H. Wang

A differentiable rule learner for inductive logic programming.

International Joint Conference on Artificial Intelligence, Jeju, Aug 2–9, 2024.

IJCAI-22 Long presentation

K. Gao, K. Inoue, Y. Cao, H. Wang

Learning First-Order Rules with Differentiable Logic Program Semantics.

International Joint Conference on Artificial Intelligence, VIENNA, July 23–29, 2022. (Virtual Attendance)

IJCLR-21 Long presentation

K. Gao, H. Wang, Y. Cao, K. Inoue

Learning from Interpretation Transition Using Differentiable Logic Semantics.

International Joint Conference on Learning & Reasoning, October 25–27, 2021. (Virtual Attendance)

Posters

AI Health Summit-23
X. Wang, J. Ong, F. Yang, Y. Xia, **K. Gao**, J. Ma, S. Goh, Y. Liu, D. Shu
Strengthening Diabetes Care Through AI-Augmented Decision Support
November 23–24, 2023, Singapore

Volunteers

AAAI-21 Student volunteer
AAAI Conference on Artificial Intelligence, February 4–7, 2021. (Virtual Attendance)

Teaching Experience

Mathematical Logic, spring semester, teaching assistant at PKU. 2019,2022
Delivered lectures; assignments and exams; F&Q.
Set Theory and Graph Theory, fall semester, teaching assistant at PKU. 2018, 2021
Assignments and exams; F&Q.
Probability and Statistics, student lecturer at USTB. Step. 2015 – June 2016
Delivered lectures and F&Q one time per two weeks.