Cai Zhou

Tsinghua University – Beijing, 100084, P.R.China
(+86)18767198389 • zhouc20@mails.tsinghua.edu.cn • Homepage

An undergraduate majoring in Automation and AI, minoring in statistics. A passionate, persistent perfectionist and a continual learner. Understand the world through theory, change the world with science and technology.

Education Background

- **Tsinghua University** — **Beijing**
  Undergraduate in Department of Automation, Class of General Artificial Intelligence — *2020-2024*
- **Tsinghua University** — **Beijing**
  Undergraduate in Statistics, Joint Degree — *2021-2024*
- **GPA: 3.93/4.0, Ranking: Top 2%**

Research Interest

My research interest lies broadly in **theoretical and applied machine learning**, with a focus on learning for (semi-)structured, discrete and combinatorial data including **graph, geometry and topology**. Currently, I aim to understand the foundation of machine learning (information theory, optimization, statistics, causal inference) and graph learning (theoretical expressive power, representation and generation, scalability). I'm also experienced in application areas including Computer Vision, Natural Language Processing and Computational Biology.

Centered at machine learning and graph learning, my long term research goal is to empower science and AI with learning across various subfields including vision, language and structured data. Concretely, I will be dedicated to the following topics: (1) **multi-modal foundation models**, or more broadly, AI systems that can perceive and reason like human; (2) **AI4science**, with an emphasis on structural biology, drug discovery and healthcare.

Publications and Research Experience

Conference Papers

- Facilitating Graph Neural Networks with Random Walk on Simplicial Complexes
  Cai Zhou, Xiyuan Wang, Muhan Zhang.
  Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS, 2023) [PDF]

  TL;DR: Design positional and structural encodings for GNNs based on random walk on simplicial complexes, specifically EdgeRWSE and Hodge1Lap for 1-simplices (edge-level) that can provably improve expressive power and practical performance.

- From Relational Pooling to Subgraph GNNs: A Universal Framework for More Expressive Graph Neural Networks
  Cai Zhou*, Xiyuan Wang*, Muhan Zhang.
  Fortieth International Conference on Machine Learning (ICML, 2023) [PDF]

  TL;DR: Propose k, l-WL algorithm (running k-WL on a graph with l labels) and theoretically establish a strict expressive power hierarchy; Incorporate a wide range of GNNs including relational pooling and subgraph GNNs.

Preprints

- On the Theoretical Expressive Power and Design Spaces of High Order Graph Transformers
  Cai Zhou, Rose Yu, Yusu Wang
  Twenty-seventh International Conference on Artificial Intelligence and Statistics (AISTATS, 2024), Under Review
  Review Scores (pre-rebuttal): 4 (accept), 4 (accept), 3 (borderline accept)

  TL;DR: Theoretically analyze the expressive power and approximation power of high order graph transformers; propose scalable and powerful high order graph transformers and simplicial transformers.

- Locally Supervised Deep Learning by Maximizing Information Propagation
  Yulin Wang, Zanlin Ni, Yifan Pu, **Cai Zhou**, Shiji Song, Gao Huang
  International Journal of Computer Vision (IJCV), Under Review

  TL;DR: Apply information theory to model locally supervised learning, propose InfoPro loss to alleviate information collapse in locally trained deep networks and verify the effectiveness in computer vision tasks.

Ongoing Projects

- Motif-based Few-shot Learning for Molecular Generation
  To be submitted to ICML 2024

  Motif-based molecular generative method that enables few-shot and in-context learning.

- Generative Graph Foundation Model
  To be submitted to ICML 2024

  Generative graph foundation model utilizing latent graph diffusion with equivariant transformer. Solve tasks of different levels (graph/node/edge), types (generation/regression/classification) and domain.

- Project: Dynamic Graph of Thought Prompting for LLM Reasoning

  To be submitted to ICML 2024

  Introduce probability and causality into LLM reasoning steps. Extend Graph of Thought (GoT) prompting to Dynamic Graph of Thought, which can model reasoning procedure dynamic and adaptive.

Academic Services

- Reviewer for AISTATS 2024
- Teaching Assistant for General Artificial Intelligence System Practice, 2023-Summer, Tsinghua University

Crucial Honors and Awards

Comprehensive Awards

- National Scholarship (Highest honor for undergraduates in China), 2023
- Comprehensive Excellence Award of Tsinghua University (Highest honor in THU and the Dept. Automation), 2023 & 2021
- Research and Innovation Excellence Award of Tsinghua University, 2022
- Academic Excellence Award of Tsinghua University, 2022
- Several prizes in national English speech contests and proficiency competitions.

Mathematics and Physics

- First prize of National Physics Competition for Undergraduates, 2021

Technical and General Skills

Technical Skills

- **Mathematics:** Calculus, Algebra, Discrete and Combinatorial Mathematics, Differential Geometry, Algebraic Topology, Complex Analysis, Fourier Analysis, Functional Analysis, Operation Research, etc.
- **Statistics:** Probability Theory, Statistical Inference, Stochastic Process, Multivariate Statistics and Regression, Computational Statistics, Time Series Analysis, Causal Inference, Bayesian Statistics, Biostatistics, etc.
- **Programming Skills:** Proficient in Python, Pytorch, PyG, DGL, C, C++, R, Linux, Latex, Markdown, Git, etc. Familiar with Tensorflow, Keras, C#, HTML, MATLAB, etc.

General Skills

Fluent in English; Good writing and presentation skills; Outstanding leadership and cooperation skills.