Guoxin Chen

(+86)18358004349 gx.chen.chn@gmail.com chen-gx.github.io

Education

Institute of Computing Technology, Chinese Academy of Sciences **Sep. 2022 – Jul. 2025**
Master of Engineering, Key Laboratory of AI Safety, Beijing, China *GPA: 3.85/4*

Zhejiang University of Technology **Sep. 2018 – Jun. 2022**
Bachelor of Engineering, College of Science, Hangzhou, China *Ranking: 1/38*

Research Experience

AlphaMath Almost Zero: Process Supervision Without Process **Mar. 2024 – Jun. 2024**
Research Intern at Institute for Intelligent Computing, Alibaba Group. *Beijing, China*

- **Motivation**: Existing approaches to mathematical reasoning primarily rely on artificially injecting external knowledge into LLMs by fine-tuning on large amounts of high-quality, process-supervised data that are heavily annotated by humans or GPT-4. These approaches overlook the reservoir of knowledge inherent in pre-trained LLMs.
- **Method**:
  1. We propose the *AlphaMath* framework, which aims to unleash the potential of **pre-trained models** without relying on GPT-4 or manual annotations, and **autonomously** enhance mathematical reasoning capabilities.
  2. *AlphaMath* leverages Monte Carlo Tree Search (MCTS) to automatically generate both **process supervision data** and **step-level evaluation signals**, bypassing the need for manual or GPT-4 annotation
  3. We propose an efficient inference strategy: *step-level beam search*, where the value model is crafted to assist the policy model (LLM) in navigating more effective reasoning paths, rather than solely relying on prior probability.
- NeurIPS 2024 Under Review (First Author). Paper Link.

Step-level Value Preference Optimization for Mathematical Reasoning **Apr. 2024 – Jun. 2024**
Research Intern at Institute for Intelligent Computing, Alibaba Group. *Beijing, China*

- **Motivation**:
  1. Supervised fine-tuning (SFT) methods only focus on positive examples, making the model blindly imitates successful cases without understanding **what the wrong solutions are**.
  2. Preference learning (DPO) can distinguish between positive and negative examples, but most efforts only focus on coarse solution-level preferences without informing **which step** in negative examples (y_l) **led to the mistake**.
  3. Although DPO simplifies the reward model in RLHF, it also **discards the value model**, which is used to estimate the expected return of the current state and can significantly improve the reasoning ability of the policy model.
- **Method**:
  1. We propose Step-level Value Preference Optimization (SVPO) to focus on the more fine-grained **step-level preference relationships** and integrate **a lightweight value model** into DPO to improve reasoning ability.
  2. We use MCTS to autonomously annotate step-level preferences, where the Q **value of each node reflects which steps may lead to mistakes in** y_l. Compared with the preferences annotated by GPT-4, this approach is more likely to highlight the mistakes that current models are prone to make.
- EMNLP 2024 Under Review (First Author). Paper Link.

Facilitating Structured Reasoning and Explanation via Reinforcement Learning **Jul. 2023 – Feb. 2024**
Research Intern at Shanghai AI Lab. *Shanghai, China*

- **Motivation**:
  1. Current QA explainable systems only furnish brief supporting evidence, without clarifying the reasoning process from premise to the derived answer.
  2. SFT methods decompose structured reasoning into single steps, **ignoring dependencies between different steps**.
  3. RL-based methods defines the return using the standard chain structure, **lacking the ability to present the tree or graph logical structures**, which hinders the potential of RL in structured reasoning.
- **Method**:
  1. We propose SEER, a RL-based method to facilitate structured reasoning and explanation, which is the first general framework that **accommodates chained, tree/graph-based** structured reasoning scenarios.
  2. We propose the **structure-based return** to define the intricate interdependencies among different reasoning steps, effectively stimulating the potential of reinforcement learning for structured reasoning.
- Accepted at ACL 2024 (First Author). Paper Link.

Multi-level Prompt Tuning for Machine Reading Comprehension **Feb. 2023 – Jun. 2023**
Research Intern at MeetYou AI Lab. *Beijing, China*

- **Motivation**: Existing soft-prompt tuning methods focus on designing **input-independent prompt vectors** for a given task, which under-utilizes the input semantics for the answer generation in machine reading comprehension.

- **Method**:
  1. We propose a **multi-level prompt tuning** (MPrompt) approach for machine reading comprehension which generates prompts at task-specific, domain-specific, and context-specific levels to enhance reasoning capabilities.
  2. We propose an **independence constraint** to steer each domain-specific prompt to focus on intra-domain information, avoid information redundancy, and enrich domain-related semantics.
  3. Proposed a **prompt generator** based on a small-scale PLM to integrate context-related knowledge into prompt generation, which enriches the semantics of generated prompt vector.
- Accepted at EMNLP 2023 (First Author) Paper Link.

Causality and Independence Enhancement for Biased Node Classification **Oct. 2022 – Feb. 2023**

Key Laboratory of AI Safety, Institute of Computing Technology. *Beijing, China*

- **Motivation**: (1) Previous methods mainly focus on a single type of data bias, such as label selection bias or structural bias; (2) Targeting only a specific bias may not necessarily improve the overall generalization performance; some methods even sacrifice the generalization of other types of data biases when improving the generalization of the specific bias.
- **Method**:
  1. the first study to analyze the impact of **mixed biases** in node classification.
  2. From the perspective of causal learning, we analyze the reasons for the poor performance of GNN in out-of-distribution generalization, and proposes CIE framework to enhance the generalization of various GNNs.
  3. CIE improves the discriminability of causal and spurious features in complex bias environments through independence constraints and mitigates spurious correlations through the backdoor adjustment.
- Accepted at CIKM 2023 (First Author) Paper Link.

Publications

Published

- **Guoxin Chen**, Kexin Tang, Chao Yang, Fuying Ye, Yu Qiao, Yiming Qian. 2023. SEER: Facilitating Structured Reasoning and Explanation via Reinforcement Learning. In *Proceedings of the 62st Annual Meeting of the Association for Computational Linguistics* (ACL 2024).
- **Guoxin Chen**, Yiming Qian, Bowen Wang, and Liangzhi Li. 2023. MPrompt: Exploring Multi-level Prompt Tuning for Machine Reading Comprehension. In *Findings of the Association for Computational Linguistics: EMNLP 2023*.
- **Guoxin Chen**, Fangda Guo, Yongqing Wang, Yanghao Liu, Peiying Yu, Huawei Shen, Xueqi Cheng. FCS-HGNN: Flexible Multi-type Community Search in Heterogeneous Information Networks. In *Proceedings of the 33nd ACM International Conference on Information and Knowledge Management.* (CIKM 2024).
- **Guoxin Chen**, Yongqing Wang, Fangda Guo, Qinglang Guo, Jiangli Shao, Huawei Shen, and Xueqi Cheng. 2023. Causality and Independence Enhancement for Biased Node Classification. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management* (CIKM 2023).

Under Review

- **Guoxin Chen**, Minpeng Liao, Chengxi Li, Kai Fan. AlphaMath Almost Zero: process Supervision without process. Submitted to NeurIPS 2024.
- **Guoxin Chen**, Minpeng Liao, Chengxi Li, Kai Fan. Step-level Value Preference Optimization for Mathematical Reasoning. Submitted to EMNLP 2024.

Major Awards & Honors

- E-Funds Fintech Freshman Fellowship awarded by ICT, Chinese Academy of Sciences Jan. 2023
- National 2nd Prize, Post-Graduate Mathematical Modeling Contest of China Nov. 2023
- E-Funds Fintech Fellowship awarded by ICT, Chinese Academy of Sciences Feb. 2024