Qinyuan Cheng

Mobile: (+86)15989016752 · Email: chengqy2019@foxmail.com
2005 Songhu Rd., Shanghai, China 200438
Personal Website · https://xiami2019.github.io/

EDUCATION

Fudan University, Computer Science, *PhD* 2021.09 - 2026.06

- **Advisor**: Prof. Xipeng Qiu
- **Research Interest**: LLM post-training: including AI alignment, reinforcement learning for LLMs , multi-modal, Agent systems.

Sun Yat-sen University, Computer Science, *Undergraduate* 2016.09 - 2020.06

- **Advisor**: Prof. Hanjiang Lai
- **Research Interest**: Unsupervised Machine Translation, Image and Cross-modal Retrieval

INTERNSHIP

Shanghai AI Laboratory, Shanghai, China 2023.05 - Now

- LLM Research Intern

ByteDance, Beijing, China 2021.03 - 2021.08

- SDE Intern (iOS app development)

PROJECTS

MOSS 2023.02 - 2023.06

- MOSS is a conversational language model like ChatGPT. It is capable of following users' instructions to perform various natural language tasks including tool using, inner thinking, question answering, generating code, etc.
- Mainly responsible for data synthesis in the post-training stage, including multi-turn dialogue, safety alignment, honest alignment.
- https://github.com/OpenMOSS/MOSS, 12k stars.

MOSS-TTSD 2025.2 - 2025.06

- MOSS-TTSD is a text to spoken dialogue generation model that enables expressive dialogue speech synthesis in both Chinese and English, supporting zero-shot multi-speaker voice cloning, voice event control, and long-form speech generation.
- I lead the whole team.
- https://github.com/OpenMOSS/MOSS-TTSD

SpeechGPT-2.0-preview 2024.10 - 2025.01

- SpeechGPT-2.0-preview a human-like real-time spoken dialogue model, meticulously trained on millions of hours of Chinese speech data. It delivers low-latency, natural responses enriched with human-like expressions, seamlessly handling interruptions and offering multi-style emotional control.
- I mainly responsible for text data curation, including pre-training and post-training.
- https://github.com/OpenMOSS/SpeechGPT-2.0-preview

Magic LLMs, HONOR 2023.08 - 2024.01

- I led a team to develop a LLM-based agent system in collaboration with HONOR, which has been integrated into Magic OS 8.0 to help users create videos from their photo albums.

Pazhou Algorithm Competition 2023.06 - 2023.12

- I organized a LLM post-training competition about improving Chinese LLMs' intelligence, truthfulness, and safety (total prize of 1,000,000 RMB).
- https://iacc.pazhoulab-huangpu.com/contestdetail?id=666abc727ff47da8cc804856&award=1,000,000

Highlighted Papers

Can AI Assistants Know What They Don't Know?

- **Qinyuan Cheng***, Tianxiang Sun*, Xiangyang Liu, Wenwei Zhang, Zhangyue Yin, Shimin Li, Linyang Li, Zhengfu He, Kai Chen, Xipeng Qiu
- ICML 2024
- We ask the question "Can AI assistants know what they don't know and express them through natural language?" To answer this question, we construct a model-specific "I don't know" (Idk) dataset for an assistant, which contains its known and unknown questions, based on existing open-domain question answering datasets. Then we design a series of truthfulness alignment methods to align the assistant with its corresponding Idk dataset and observe whether it can refuse to answer its unknown questions after alignment.

Evaluating Hallucinations in Chinese Large Language Models

- **Qinyuan Cheng**, Tianxiang Sun, Wenwei Zhang, Siyin Wang, Xiangyang Liu, Mozhi Zhang, Junliang He, Mianqiu Huang, Zhangyue Yin, Kai Chen, Xipeng Qiu
- Preprint, Cited by OpenAI and DeepMind
- We introduce HalluQA, a benchmark with 450 adversarial questions to assess hallucinations in Chinese large language models, covering diverse domains and cultural contexts. It targets imitative falsehoods and factual errors, built using GLM-130B and ChatGPT, and evaluated automatically with GPT-4.

Survey Paper: Scaling of Search and Learning: A Roadmap to Reproduce o1 from Reinforcement Learning Perspective

- **Qinyuan Cheng***, Zhiyuan Zeng*, Zhangyue Yin*, Bo Wang*, Shimin Li, Yunhua Zhou, Qipeng Guo, Xuanjing Huang, Xipeng Qiu
- Attract widespread attention in the global AI community.
- We analyze OpenAI's o1, an AI milestone excelling in reasoning tasks, driven by reinforcement learning. Unlike alternatives like knowledge distillation, limited by teacher models, our roadmap focuses on four key components: policy initialization for human-like reasoning, reward design for effective guidance, search for high-quality solutions, and learning to enhance performance with more data and parameters. These elements highlight how learning and search power o1's success, influencing LLM development

Improving Contrastive Learning of Sentence Embeddings from AI Feedback

- **Qinyuan Cheng**, Xiaogui Yang, Tianxiang Sun, Linyang Li, Xipeng Qiu
- Findings of ACL 2023
- We propose CLAIF and CLHAIF, the first work to use AI feedback to provide scalable training signals for Contrastive Learning and combine AI feedback with human feedback for futher improvement. Unlike typical methods struggling with sample pair quality due to language's discrete nature, CLAIF uses AI feedback from large language models to create fine-grained similarity scores for sample pairs.

Unified Active Retrieval for Retrieval Augmented Generation

- **Qinyuan Cheng***, Xiaonan Li*, Shimin Li, Qin Zhu, Zhangyue Yin, Yunfan Shao, Linyang Li, Tianxiang Sun, Hang Yan, Xipeng Qiu
- Findings of EMNLP 2024
- We introduce Unified Active Retrieval (UAR) to improve Retrieval-Augmented Generation (RAG) by addressing inefficiencies in existing active retrieval methods. UAR uses four orthogonal criteria as simple classification tasks for better retrieval timing, with minimal extra cost. Supported by UAR-Criteria, it handles diverse instructions effectively.

Is MultiWOZ a Solved Task? An Interactive TOD Evaluation Framework with User Simulator

- **Qinyuan Cheng***, Linyang Li*, Guofeng Quan, Feng Gao, Xiaofeng Mou, Xipeng Qiu
- Findings of EMNLP 2022
- We construct an User Simulator for task-oriented agents and use SFT + RL to fine-tune t5 models. Based on these, we propose an interactive evaluation framework for TOD agents and use reward-shaping to improve the quality of the generated responses.

How to Mitigate Overfitting in Weak-to-strong Generalization?

- Junhao Shi*, **Qinyuan Cheng***, Zhaoye Fei, Yining Zheng, Qipeng Guo, Xuanjing Huang, Xipeng Qiu

- ACL 2025
- In this paper, we investigate how to mitigate overfitting in weak-to-strong generalization. We propose a two-stage training framework to both control the label quality and questions difficulties. Our framework achieve superior performance compared with other weak-to-strong baselines.

Revisiting the Test-Time Scaling of o1-like Models: Do they Truly Possess Test-Time Scaling Capabilities?

- Zhiyuan Zeng, **Qinyuan Cheng**, Zhangyue Yin, Yunhua Zhou, Xipeng Qiu
- ACL 2025
- The test-time scaling ability is crucial for o1-like models. However, we find that many models do not have true test-time scaling abilities. Besides, we propose a test-time sampling methods to enhance models' test-time scaling abilities.x

ALL PUBLICATIONS

- **Mitigating Negative Style Transfer in Hybrid Dialogue System**
  Shimin Li, **Qinyuan Cheng**, Linyang Li, Xipeng Qiu
  AAAI 2023
- **Inference-Time Decontamination: Reusing Leaked Benchmarks for Large Language Model Evaluation**
  Qin Zhu, **Qingyuan Cheng**, Runyu Peng, Xiaonan Li, Tengxiao Liu, Ru Peng, Xipeng Qiu, Xuanjing Huang
  EMNLP 2024 (Findings)
- **Cross-Modality Safety Alignment**
  Siyin Wang, Xingsong Ye, **Qinyuan Cheng**, Junwen Duan, Shimin Li, Jinlan Fu, Xipeng Qiu, Xuanjing Huang
  NAACL 2025 (Findings)
- **Scaling Laws for Fact Memorization of Large Language Models**
  Xingyu Lu, Xiaonan Li, **Qinyuan Cheng**, Kai Ding, Xuanjing Huang, Xipeng Qiu
  EMNLP 2024 (Findings), Cited by DeepMind
- **World Modeling Makes a Better Planner: Dual Preference Optimization for Embodied Task Planning**
  Siyin Wang, Zhaoye Fei, **Qinyuan Cheng**, Shiduo Zhang, Panpan Cai, Jinlan Fu, Xipeng Qiu
  ACL 2025
- **VisuoThink: Empowering LVLM Reasoning with Multimodal Tree Search**
  Yikun Wang, Siyin Wang, **Qinyuan Cheng**, Zhaoye Fei, Liang Ding, Qipeng Guo, Dacheng Tao, Xipeng Qiu
  ACL 2025
- **Agent alignment in evolving social norms**
  Shimin Li, Tianxiang Sun, **Qinyuan Cheng**, Xipeng Qiu
  arXiv:2401.04620
- **Perceive the Passage of Time: A Systematic Evaluation of Large Language Model in Temporal Relativity**
  Shuang Chen, Yining Zheng, Shimin Li, **Qinyuan Cheng**, Xipeng Qiu
  COLING 2025
- **MOSS: An Open Conversational Large Language Model**
  Tianxiang Sun, Xiaotian Zhang, Zhengfu He, Peng Li, **Qinyuan Cheng**, Xiangyang Liu, Hang Yan, Yunfan Shao, Qiong Tang, Shiduo Zhang, Xingjian Zhao, Ke Chen, Yining Zheng, Zhejian Zhou, Ruixiao Li, Jun Zhan, Yunhua Zhou, Linyang Li, Xiaogui Yang, Lingling Wu, Zhangyue Yin, Xuanjing Huang, Yu-Gang Jiang, Xipeng Qiu
  Machine Intelligence Research
- **LLM can Achieve Self-Regulation via Hyperparameter Aware Generation**
  Siyin Wang, Shimin Li, Tianxiang Sun, Jinlan Fu, **Qinyuan Cheng**, Jiasheng Ye, Junjie Ye, Xipeng Qiu, Xuanjing Huang
  ACL 2024 (Findings)
- **Explicit memory learning with expectation maximization**
  Zhangyue Yin, Qiushi Sun, Qipeng Guo, Zhiyuan Zeng, **Qinyuan Cheng**, Xipeng Qiu, Xuanjing Huang
  EMNLP 2024
- **Dictionary learning improves patch-free circuit discovery in mechanistic interpretability: A case study on othello-gpt**

Zhengfu He, Xuyang Ge, Qiong Tang, Tianxiang Sun, **Qinyuan Cheng**, Xipeng Qiu
arXiv:2402.12201 2024, Cited By Anthropic

- **In-Memory Learning: A Declarative Learning Framework for Large Language Models**
  Bo Wang, Tianxiang Sun, Hang Yan, Siyin Wang, **Qingyuan Cheng**, Xipeng Qiu
  arXiv:2403.02757 2024
- **Case2Code: Scalable Synthetic Data for Code Generation**
  Yunfan Shao, Linyang Li, Yichuan Ma, Peiji Li, Demin Song, **Qinyuan Cheng**, Shimin Li, Xiaonan Li, Pengyu Wang, Qipeng Guo, Hang Yan, Xipeng Qiu, Xuan-Jing Huang, Dahua Lin
  COLING 2025
- **Reasoning in flux: Enhancing large language models reasoning through uncertainty-aware adaptive guidance**
  Zhangyue Yin, Qiushi Sun, Qipeng Guo, Zhiyuan Zeng, Xiaonan Li, Junqi Dai, **Qinyuan Cheng**, Xuanjing Huang, Xipeng Qiu
  ACL 2024
- **Aggregation of Reasoning: A Hierarchical Framework for Enhancing Answer Selection in Large Language Models**
  Zhangyue Yin, Qiushi Sun, Qipeng Guo, Zhiyuan Zeng, Xiaonan Li, Tianxiang Sun, Cheng Chang, **Qinyuan Cheng**, Ding Wang, Xiaofeng Mou, Xipeng Qiu, XuanJing Huang
  COLING 2024

ACADEMIC SERVICES

Conference Reviewer / Program Committee Member

- ACL (2023, 2024, 2025)
- Neurips (2024, 2025)
- ICLR (2025)
- ICML (2025)
- AISTATS (2025)
- EMNLP (2022, 2023, 2024, 2025)
- COLING (2025)
- AISTATS (2025)