

CHUYAN ZHOU

✉ chyzhou24@gmail.com · 🌐 Github · 🏠 Homepage

EDUCATION

Institute of Science Tokyo, Tokyo, Japan 2026.9 – 2028.9 (expected)

Master Degree in Artificial Intelligence, School of Computing

Advised by Prof. Naoaki Okazaki

University of California, Berkeley, California, USA 2024.8 – 2025.1

GLOBE Program in Computer Science major, College of Engineering

Overall GPA 4.0/4.0

ShanghaiTech University, Shanghai, China 2022.9 – 2026.7

Bachelor of Engineering in Computer Science

Advised by Prof. Kewei Tu

Overall GPA 3.85/4.0, ranked 3/162 (top 2%), *Major GPA* 4.0/4.0

RESEARCH EXPERIENCE & PAPERS UNDER REVIEW

Okazaki Laboratory Inst. of Science Tokyo 2026.9 – Present

Some of my research interests for the time being:

1. Diffusion LLMs: their architecture, the forward-denoising (inference) process, their agentic RL training
2. Agentic RL and other agentic training paradigms
3. Recursive Self-improvement of LLMs

VDI Center ShanghaiTech University 2023.10 – 2026.7

Research Intern Advised by Prof. Kewei Tu

Training-free Verifiable Process Reward for LLM Reinforcement Finetuning Archived
Contributor

- In modern LLM RL finetuning, ORMs are common of rewarding scheme to predict by the final outcome of the task. Though simple, ORMs suffer reward sparsity causing LLMs' failure to generate useful thinking process, leading to possibly degenerate performance. PRMs are to predict the reward by intermediate steps of the LLMs, providing more fine-grained/informative feedback.
- However, existing PRMs either (1) are black-box neural models, requiring human annotations & training; (2) are training-free/nonparametric by statistical heuristics, while trajectories are segmented into semantically incomplete units; (3) are aggregated into some outcome reward for each trajectory, not really doing credit assignment.
- We designed a training-free verifiable process rewarding scheme based on formalized theorem provers, consisting of:
 - Forcing LLMs to generating natural language reasoning steps in Horn clause $p \wedge q \rightarrow r$ form, enabling semantically complete segmentation of the reasoning process;
 - A training-free process reward by integrating a theorem prover/solver called a verifier;
 - An LLM teacher in-the-loop to formalize NL reasoning into verifiers;
 - A GRPO-like advantage normalizing algorithm for critic-free value estimation, but for PRMs.
- We base experiments on logical/math QA datasets to train reasoning models, and evaluate and compare the performance by accuracy improvement after RL finetuning with our proposed process reward.

Augmenting Transformer Language Models with Dependency Graphs ACL 2026 Main Conference
Author

- In this project, we proposed a novel method to augment Transformer-based syntactic language models with dependency graphs, especially with word-synchronous semantic dependency graphs, which is to model

the dependencies and tokens jointly in a unified framework with a dependency-predicting head so that the original word sequences are kept intact.

- Since the use of syntactic information in past works is limited to recursive structures i.e. trees, and graphs are more general structures which we consider have more potential information to be utilized, based on a codebase of Transformer-XL, we implemented the model architecture which operates on the dependency graphs and word sequences jointly.
- As a contributor, I conducted baseline experiments on metrics such as perplexity (PTB 1987-89 WSJ Treebanks), SG and BLiMP tests.

MISCELLANEOUS PROJECTS & COMPETITIONS

LLM-powered Lecture Generation Application UC Berkeley, LLM Agent Hackathon 2024.8 – 2024.12
Team Leader Project, Team of 5

- Independently developed the backend framework of the lecture generation pipeline based on FastAPI as a deployable web service. The backend service involves asynchronous task running using multithreading, task managing by API powered by Redis databases, and a metadata system managing the generated data.
- Worked as a main developer and integrated the respective model components, which run on the above backend framework.
- Developed an additional LLM-powered QA agent based on the built backend, which involves interacting with LLMs using a long context of generated lectures with a RAG system dynamically indexing the sources (e.g. textbooks) for LLM grounding.

Multi-step ASR Pipeline Bengali.AI Speech Recognition Challenge, Kaggle 2023.7 – 2023.10
Team Leader Competition, Team of 2

- We developed an automatic speech recognition (ASR) pipeline for a low-resource language (Bangla/Bengali) in a team of 2. With supervised finetuning from augmented low-resource speech datasets, the inference pipeline mainly consists of denoising → inference main ASR decoder model using finetuned Wav2Vec2 whose decoder probabilities are fused with those of an n-gram draft language model → post-processing by rearranging the punctuation from plain words using a BERT-based punctuation prediction model.
- I worked mainly on the n-gram & punctuation enhancing and the finetuning of the Wav2Vec2 model.

ACTIVITIES

SI 140A Course (Probability and Statistics for EECS) ShanghaiTech 2025.2 – 2025.7
Head TA

- Held discussion classes in a weekly basis for 40 students, covering the course materials, assignments and providing additional insights. Mainly responsible for creating & grading assignments and exams.

HONORS AND AWARDS

- ***Distinguished Alumni Award***, ShanghaiTech University 2026.7
- ***Outstanding Student Award & Scholarship***, ShanghaiTech University 2025.5
- ***Silver Medal*** (24-th place), Award on Bengali.AI Speech Recognition Challenge on Kaggle 2023.10