Qingyang Zhu

☐ +86 135 1288 4971 | @ zhuqingyang.zqy@gmail.com | ◉ https://martianmartina.github.io

EDUCATION

ShanghaiTech University   Shanghai, China
B.Eng. in Computer Science and Technology;   *Sep 2019 – Jun 2023*
*Major GPA: **3.98/4.00** | Overall GPA: **3.88/4.00** | rank in school: **1/226***

- **Relevant coursework:** Artificial Intelligence (A+), Introduction to Machine Learning (A+), Numerical Optimization (A), Operating Systems (A+), Introduction to Algorithmic Game Theory (A), Algorithms and Data Structures (A+), Computer Architecture (A+)

Massachusetts Institute of Technology   Cambridge, United States
*Undergraduate Special Student; GPA: **5.0/5.0***   *Jan 2022 – Dec 2022*

- **Relevant coursework:** Matrix Methods (A+), Quantitative Methods for NLP (A), Advances in Computer Vision (A), Design and Analysis of Algorithms (A), Interactive Music Systems (A), Undergraduate Research (P)

PUBLICATIONS

Hu, Xiang, Qingyang Zhu, Kewei Tu, and Wei Wu. *Augmenting transformers with recursively composed multi-grained representations.* International Conference on Learning Representations (ICLR). Vienna, Austria. May 2024 (To Appear). arXiv: 2309.16319 `[cs.CL]`.

RESEARCH EXPERIENCE

Recursive Hierarchical Language Modeling   Shanghai, China
Intern | Ant Group   *April 2023 – present*
Research Assistant | ShanghaiTech University   *July 2023 – present*

- Advisors: Xiang Hu, Kewei Tu
- Worked on a novel recursive encoding layer capable of learning unsupervised contextualized constituent representations in logarithmic time, which can be jointly pretrained with Transformers at a large scale.
- Designed and conducted thorough downstream experiments on GLUE and span-level tasks. Achieved performance comparable with BERT's on span-level tasks with a much smaller dataset.
- Adapted the model to be a novel tokenizer that leverages the structure provided by the encoder and provides segmentation that aligns better with morphological knowledge. (ongoing new project)
- Picked up several training techniques such as Distributed Data Parallelism training, mixed-precision training, optimizer warm-up, gradient scaling, and clipping for efficient and better optimized code.

Neural QCFG with Heuristic Phrase Alignment   Shanghai, China
Undergraduate Thesis | ShanghaiTech University   *Oct 2022 – June 2023*

- Advisor: Kewei Tu
- Improved neural *quasi-synchronous context-free grammar* (QCFG) on machine translation using posterior regularization on a heuristic alignment score, which leveraged a pretrained multilingual language model.
- Implemented the Posterior Regularization by applying golden-section search to solve the dual problem and performing algorithms on structured models with a hacked PyTorch-struct framework.
- Further adapted and evaluated the method on GEO-Aligned dataset (GeoQuery with annotated word alignments). Our method can even surpass vanilla QCFG trained with gold word alignments.

Reward-conditioned Sequence Generation   Cambridge, MA, United States
Undergraduate Researcher | CSAIL, MIT   *July 2022 – Jan 2023*

- Advisor: Yoon Kim
- Enhancing machine translation through sequence-level and word-level reward conditioning.
- Designed continuous and discrete rewards of multiple granularities indicating the BLEU score level and conducted thorough experiments on reward representations and training objectives.
- Developed all the code based on Facebook's fairseq framework by hacking the architecture and building new user-defined modules.

Teaching Experience

Teaching Assistant   Shanghai, China

Discrete Mathematics   *Feb 2021 − Jun 2021*

Algorithms and Data Structures   *Sept 2021 − Jan 2022*

- Developing and grading homework and exams.
- Holding a weekly recitation class to help students review the course content.
- Staffing weekly office hours to answer technical questions and help with debugging.
- Awarded as "**2021 SIST Outstanding TA**"

Honors & Awards

First-prize scholarship for exchange programs | ShanghaiTech University

Recognition of Distinction | Drexel University

Attended Brain Technology Convergence global online classroom at Drexel University in summer 2022.

Completed with **top 5%** performance and was invited as a **peer mentor** for summer 2023.

SIST Outstanding TA 2021 | ShanghaiTech University

Outstanding Award of Women in Technology Hackathon | XILINX

Built a self-driving car using a XILINX PYNQ board. Ranked 4th in the race competition.

Outstanding Student 2020− 2021 | ShanghaiTech University: Top 2% overall performance.

Outstanding Student 2019− 2020 | ShanghaiTech University: Top 3% − top 8% overall performance

National English Competition for College Students 2021, Second Prize for Band C

Course Projects

Controllable Face Synthesis via NMFed latent space | Matrix Methods, Advances in Computer Science

- Experimented with performing Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NMF) on the latent space of an Autoencoder to find interpretable and disentangled (only for NMF) features to control the image synthesis.

Classical Music Generator | Artificial Intelligence, Introduction to Machine Learning

- A classical music generator that can extend a given melody using Markov Chain (MC) and generate the accompaniment using Hidden Makov Models (HMM). Also explored the relationship between note attributes (pitch and rhythm) by replacing MC with Dynamic Bayesian Network (DBN). Further experimented with a LSTM.

PintOS | Operating System

- Implemented Thread Scheduler, User Programs, Virtual Memory and File Systems. The same operating system in Stanford CS140.

Sing Go Hero | Interactive Music System

- A karaoke platformer game where players use the pitch of their voice to control the avatar, with the added bonus of tempo change and voice filter shenanigans. Implemented with Kivy framework.

Selling to A No-Regret Buyer | Introduction to Algorithmic Game Theory

- A project analyzing the best seller strategy when facing with a specific no-regret buyer by running simulation with different buyer algorithms.

Skills

Programming: C, C++, Python, MATLAB

Frameworks: NumPy, PyTorch, fairseq, HuggingFace, Keras, Pandas, scikit-learn

Languages: Chinese (Native), English (Fluent), French (Intermediate), Japanese (Elementary)

TOEFL score: 119 (Reading: 30, Listening: 30, Speaking: 30, Writing: 29)