

# Ruizhe Shi

 [ruizshi03@gmail.com](mailto:ruizshi03@gmail.com)  [srzer](#)

## Education

### Tsinghua University, Beijing, China

Sept. 2021 – Present

*Undergraduate at Special Pilot Class in Computer Science (Yao class)*

- Major: *Computer Science and Technology*
- Minor: *Chinese Language and Literature*

GPA 3.93/4.00

GPA 4.00/4.00

### University of Washington, Seattle, USA

Feb. 2024 – July 2024

*Research Visitor at Paul G. Allen School of Computer Science*

## Publications

(\* indicates equal contribution.)

[4] Ruizhe Shi, Yifang Chen, Yushi Hu, Alisa Liu, Hannaneh Hajishirzi, Noah A. Smith, Simon S. Du.

Decoding-Time Language Model Alignment with Multiple Objectives.[\[link\]](#)

*ICML Workshop on Theoretical Foundations of Foundation Models*, 2024.

[3] Chenhao Lu, Ruizhe Shi\*, Yuyao Liu\*, Kaizhe Hu, Simon S. Du, Huazhe Xu.

Rethinking Transformers in Solving POMDPs.[\[link\]](#)

*International Conference on Machine Learning*, 2024.

[2] Ruizhe Shi\*, Yuyao Liu\*, Yanjie Ze, Simon S. Du, Huazhe Xu.

Unleashing the Power of Pre-trained Language Models for Offline Reinforcement Learning.[\[link\]](#)

*International Conference on Learning Representations*, 2024.

[1] Yanjie Ze, Yuyao Liu\*, Ruizhe Shi\*, Jiaxin Qin, Zhecheng Yuan, Jiashun Wang, Huazhe Xu.

H-InDex: Visual Reinforcement Learning with Hand-Informed Representations for Dexterous Manipulation.[\[link\]](#)

*Annual Conference on Neural Information Processing Systems*, 2023.

## Research Experience

### DPO Theory

May 2024 – Present

*Supervised by Prof. Simon S. Du*

*CSE, University of Washington*

- In preparation for **ICLR 2025**.

### Multi-Objective Language Model Alignment

Dec. 2023 – May 2024

*Supervised by Prof. Simon S. Du*

*CSE, University of Washington*

- We propose a training-free, simple yet effective decoding-time algorithm for multi-objective alignment of language models, with optimality guarantees. Under review.
- We exploit a common form among a family of  $f$ -divergence regularized alignment approaches (such as PPO, DPO, and their variants) to identify a closed-form solution by Legendre transform, and derive an efficient decoding strategy. Theoretically, we show why existing approaches can be sub-optimal even in natural settings and obtain optimality guarantees for our method.

### Representation Theory of Transformer for Reinforcement Learning

Nov. 2023 – Jan. 2024

*Supervised by Prof. Huazhe Xu*

*IIIS, Tsinghua University*

- We challenge the common wisdom and prove theoretically and empirically that Transformers are not suitable for Partially Observable RL, while advocating Linear RNN as a promising alternative. Second-authored work accepted by **ICML 2024**.
- We establish that regular languages, which Transformers struggle to model, are reducible to POMDPs. This poses a significant challenge for Transformers in learning POMDP-specific inductive biases, due to their lack of inherent recurrence found in other models like RNNs, with empirical results highlighting the sub-optimal performance of the Transformer and considerable strength of LRU.

### Tuning Language Model for Offline Reinforcement Learning

June 2023 – Sept. 2023

*Supervised by Prof. Huazhe Xu*

*IIIS, Tsinghua University*

- We leverage the power of pre-trained Language Models for low-level motion control in offline reinforcement learning. First-authored work accepted by **ICLR 2024**.
- We demonstrate the superiority of LaMo over DT-based and value-based offline RL algorithms. Specifically, we find that LaMo could successfully handle the challenging low-data regime while DT could not. This highlights the great potential of our cross-domain pre-training for sequential modeling.

## Visual Representation for Reinforcement Learning

Mar. 2023 – May 2023

Supervised by Prof. Huazhe Xu

IIIS, Tsinghua University

- We propose H-InDex, a hand-informed visual representation for dexterous manipulation with reinforcement learning. Second-authored work accepted by **NeurIPS 2023**.
- We show the effectiveness of our framework on 12 challenging visual dexterous manipulation tasks, comparing with recent strong foundation models such as VC-1. Our study has offered valuable insights into the application of pre-trained models for dexterous manipulation, by exploring the direct application of a 3D human hand pose estimation model

## Awards & Honors

### Yao Award (Silver Medal)

Sept. 2024

IIIS, Tsinghua University

top scholarship; 3 students institute-wide

### Jiang Nanxiang Scholarship

Nov. 2023

Tsinghua University

top scholarship; 1 student per major

### China National Endeavor Scholarship

Oct. 2022

Beijing Education Bureau

1 student per major

### First Prize in College Student Mathematics Competition

Oct. 2022

Chinese Mathematical Society

### First Prize in National High School's Mathematics Competition of China

Oct. 2020

Chinese Mathematical Society

top 20

## Service

### Conference Reviewer

July 2024

NeurIPS 2024

Online

### Workshop Program Committee

Oct. 2023

FMDM 2023 at NeurIPS [\[link\]](#)

Online

### Teaching Assistant

Sept. 2024 – Jan. 2025(expected)

Natural Language Processing

Tsinghua University

### Drop-in Tutoring

Oct. 2022 – July 2024

STEM Courses

Tsinghua University

I have **157** hours of officially recorded volunteering work.

## Selected Courses

**Mathematics and Theory:** Calculus (**A**<sup>+</sup>), Linear Algebra (**A**), Abstract Algebra (**A**), Introduction to Complex Analysis (**A**), Probability and Statistics (**A**), Basic Topology (**93**), Introduction to Optimization (**A**), Theory of Computation (**A**), Physics of Information (**A**);

**Programming and AI:** Introduction to Programming in C/C++ (**A**<sup>+</sup>), Intelligent Unmanned System (**A**<sup>+</sup>), Type-safe Modern System Practice (**A**), Machine Learning (**A**), Artificial Intelligence: Principles and Techniques (**A**), Natural Language Processing (**A**).

## Technical Skills

**Programming Skills:** Python, C/C++, L<sup>A</sup>T<sub>E</sub>X, Bash, Scala, Matlab.

**Language Skills:** Chinese Mandarin (native), English (CET-6, TOEFL 104 [R30/L26/S23/W25]).