Runlong Zhou (周润龙)

September 26, 2024

Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, Washington, U.S.A.
zhourunlongvector@gmail.com, vectorzh@cs.washington.edu
https://github.com/zhourunlong

Research Interests

- Emphasis on reinforcement learning theory

- Applied machine learning and reinforcement learning

Education

- **University of Washington**
  PhD, Paul G. Allen School of Computer Science & Engineering, advised by Prof. Simon S. Du
  Seattle, USA
  2022.9 - Now (Est. 2027.8)

- **Tsinghua University**
  BEng, Special Pilot CS Class (Yaoclass), Institute for Interdisciplinary Information Sciences (IIIS)
  GPA: 3.84 (overall) / 3.89 (major) over scale 4.0, Rank: 16 / 54
  Beijing, China
  2018.8 - 2022.6

Academic Experience

- **Microsoft AI**
  Research intern with Dr. Yi Zhang
  Redmond, USA
  2024.6 - 2024.8

- **Microsoft Research**
  Research intern with Dr. Beibin Li
  Redmond, USA
  2023.6 - 2023.9

- **University of Washington**
  Research intern with Prof. Simon S. Du
  Virtual
  2020.9 - 2022.9

- **Facebook AI Research**
  Research intern with Dr. Alessandro Lazaric and Dr. Matteo Pirotta
  Virtual
  2021.3 - 2021.5

Publications

* denotes equal contribution or alphabetical ordering.

1. **Reflect-RL: Two-Player Online RL Fine-Tuning for LMs** [Link]
   Runlong Zhou, Simon S. Du, Beibin Li
   ACL 2024   Poster
   We developed Reflect-RL, a two-player system to align language models with interactive decision-making tasks. Techniques included are reflection, negative example generation, single-prompt action enumeration, and curriculum learning.

2. **Free from Bellman Completeness: Trajectory Stitching via Model-based Return-conditioned Supervised Learning** [Link]
   Zhaoyi Zhou, Chuning Zhu, **Runlong Zhou**, Qiwen Cui, Abhishek Gupta, Simon Shaolei Du
   ICLR 2024   Poster
   We study the strengths and weaknesses of return-conditioned supervised learning, and propose an empirically improved algorithm.

3. **Sharp Variance-Dependent Bounds in Reinforcement Learning: Best of Both Worlds in Stochastic and Deterministic Environments** [Link]
Runlong Zhou, Zihan Zhang, Simon S. Du
ICML 2023   Poster
We provide a systematic study of variance-dependent regret bounds of model-based and model-free reinforcement learning for tabular MDPs. The proposed model-based algorithm is both optimal for stochastic and deterministic MDPs.

4. **Variance-Dependent and Horizon-Free Reinforcement Learning for Latent Markov Decision Processes** [Link]
Runlong Zhou, Ruosong Wang, Simon S. Du
ICML 2023   Poster
We provide an algorithm framework for Latent MDPs (with context in hindsight), achieving the first horizon-free minimax regret. We complement the study by giving a novel regret lower bound for LMDPs using the symmetrization technique.

5. **Understanding Curriculum Learning in Policy Optimization for Online Combinatorial Optimization** [Link]
Runlong Zhou, Zelin He, Yuandong Tian, Yi Wu, Simon S. Du
TMLR
We formulate of canonical online Combinatorial Optimization problems as Latent MDPs and give convergence guarantee of Natural Policy Gradient on LMDPs. We show effectiveness of Curriculum Learning through the perspective of relative conditional number.

6. **Stochastic Shortest Path: Minimax, Parameter-Free and Towards Horizon-Free Regret** [Link]
Jean Tarbouriech*, **Runlong Zhou***, Simon S. Du, Matteo Pirotta, Michal Valko, Alessandro Lazaric
NeurIPS 2021   Spotlight, 3% acceptance rate
We propose an algorithm (EB-SSP) for SSP problems, which is the first to achieve minimax optimal regret while being parameter-free.

Preprints

* denotes equal contribution or alphabetical ordering.

1. **Multi-Agent Reinforcement Learning from Human Feedback: Data Coverage and Algorithmic Techniques** [Link]
Natalia Zhang*, Xinqi Wang*, Qiwen Cui*, **Runlong Zhou**, Sham M. Kakade, Simon S. Du
We study Multi-Agent Reinforcement Learning from Human Feedback (MARLHF) by exploring both theoretical foundations and empirical validations. Included in our proposed methods are Mean Squared Error (MSE) regularization and imitation learning.

Ongoing Projects

1. **DPO Theories**
Advised by Prof. Simon Du
We aim to give hardness and separation results of the empirically successful DPO algorithm.

2. **Physics of Language Models**
Advised by Dr. Yi Zhang
We study how language models learn and manipulate knowledge.

Awards, Grants & Honors

- **Undergraduate:**

- **Secondary school:**

Past Projects

If hyperlink not applicable, please refer to my GitHub page.

1. **RhythmicSpeechSongAligner** [Link]
   Runlong Zhou*, Zhui Zhu*
   Multimedia Computing course project   python
   Create visually-rhythmic videos by matching each lyric with a video clip which has similar semantic meaning

2. **GodScanner** [Link]
   Rui Shen*, Heyang Zhao*, **Runlong Zhou***
   Deep Learning course project   python
   Transfer a photo of file into its scanned style (flatten, shadow and watermark removal)

3. **Ray Tracing Renderer** [Link]
   Runlong Zhou
   Advanced Computer Graphics course project   C++
   Optimized path tracing framework supporting Mitsuba configurations, many textures and sampling methods

4. **Texas Hold'em Agent** [Link]
   Runlong Zhou
   Game Theory course project   C++
   A smart Texas hold'em agent

5. **IconAdapter** [Link]
   Kailu Wu*, **Runlong Zhou***
   Machine Learning course project   python
   Transfer icon styles to match mobile UI themes

Miscellanea

- **Professional skills:** Algorithm design, Data structures, Deep (Reinforcement) learning

- **Programming skills:** C++ / C, python, LaTeX, HTML, JavaScript (Node.js), CUDA, Java, MATLAB, Rust

- **Hobbies:** Air rifle / pistol shooting, Archery

- **Teaching:** Teaching Olympiad in Informatics to secondary school students between 2018 and 2021. Teaching assistant of UW CSE 543 Deep Learning, CSE 446 Machine Learning.

- **Reviewing:** ICML 2022, NeurIPS 2022, ICML 2024, COLT 2024.