SHENAO ZHANG

shenao@gatech.edu | shenao-zhang.github.io

EDUCATION

Georgia Institute of Technology *May 2020 - May. 2022*
M.S. in ECE (Electrical and Computer Engineering), GPA: 3.81/4.00 *Atlanta, GA*

South China University of Technology *Aug. 2016 - May 2020*
B.Eng. in EE (Electronic and Information Engineering, Innovation Class) *Guangzhou, China*

University of California, Berkeley *Jan. 2019 - May 2019*
Visiting student at Department of EECS, GPA: 3.90/4.00 *Berkeley, CA*

RESEARCH INTERESTS

My research centers around reinforcement learning (RL), especially model-based RL. I'm interested in developing and analyzing efficient RL algorithms with application to robotic and multi-agent systems.

PUBLICATIONS AND PREPRINTS

[1] **Shenao Zhang**, "Conservative Dual Policy Optimization for Efficient Model-Based Reinforcement Learning", *Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS)*, 2022. [PDF].

[2] **Shenao Zhang**, Boyi Liu, Yan Li, Zhaoran Wang, Tuo Zhao, "Model-Based Reparameterization Policy Gradient: Theory and Practical Algorithms", in submission to *International Conference on Machine Learning (ICML)*, 2023. [PDF].

[3] **Shenao Zhang**, Wanxin Jin, Zhaoran Wang, "Model-Based First-Order Policy Gradient for Contact Dynamics", in submission to *International Conference on Machine Learning (ICML)*, 2023. [PDF]

[4] **Shenao Zhang**, Li Shen, Lei Han, Li Shen, "Learning Meta Representation for Agents in Multi-Agent Reinforcement Learning", *ICLR Workshop on Gamification and Multiagent Solutions*, 2022. Under review at *Machine Learning Journal*. [PDF]

[5] **Shenao Zhang**, Li Shen, Zhifeng Li, Wei Liu, "Structure-Regularized Attention for Deformable Object Representation", *NeurIPS Workshop on Object Representations for Learning and Reasoning*, 2020. [PDF]

[6] Dazheng Hu, Huabiao Qin*, Hongmei Liu*, **Shenao Zhang***, "Gaze Tracking Algorithm Based on Projective Mapping Correction and Gaze Point Compensation in Natural Light", *International Conference on Control and Automation (ICCA)*, 2019. [PDF]

RESEARCH EXPERIENCE

Northwestern University *Aug. 2022 - Present*
Research Intern. Advisor: Zhaoran Wang *Remote*

- **First-Order Policy Gradient for Contact Dynamics** (in progress): We study the FOPG with complementarity-based models for robotics systems that experience hard contact. We show that the convergence and gradient variance of FOPG are closely related to the model stiffness, which is determined by the centering parameter μ when solving the complementarity problem. Therefore, we propose the Analytic Barrier Smoothing with a contact-aware centering parameter. By discovering its connection with randomized smoothing, we show that analytic smoothing is the best linear approximation of the LCP solution and establish its gradient bias upper bound.

Georgia Tech *Sep. 2020 - Aug. 2022*
Research Intern. Advisors: Tuo Zhao and Zhaoran Wang *Atlanta, GA*

- **Model-Based Reparameterization Policy Gradient Methods**: We establish the first convergence analysis result for model-based RP PGMs, and our theory identifies the smoothness of the

function approximators as a major determining factor that affects the quality of gradient estimation. Based on our theory, we further propose a spectral normalization method, which can effectively mitigate the exploding variance due to long model unrolls. Experiments are provided to support our theory and method: With a proper normalization, we can significantly reduce the gradient variance of modelbased RP PGMs and improve their convergence, leading to equal or better performance than their counterparts based on other gradient estimators, e.g., Likelihood Ratio (LR) gradient estimator.

- **Conservative Dual Policy Optimization for Efficient MBRL**: Previous provable model-based RL achieves the global optimality by assuming that the model has restricted complexity, which rarely holds in nonlinear settings. In practice when the model is poorly generalized, as measured by the model complexity, an exploration step can only eliminate a small portion of the model hypothesis, causing over-exploration. To solve this issue, I proposed *Conservative Dual Policy Optimization* (CDPO). CDPO optimizes an intermediate policy under a stable *reference* model and then maximizes the *expected* policy value. I proved the statistical equivalence between CDPO and posterior-sampling RL in Bayesian regret, with which the global optimality of CDPO can be concluded. Moreover, I established the monotonic policy value improvement result of CDPO to ensure efficient exploration. The experiments in tabular and MuJoCo tasks support the claims and reveal the superiority of the proposed method.

Tencent AI Lab *Aug. 2019 - Sep. 2020*
Research Intern. Advisors: Li Shen, Lei Han and Li Shen *Shenzhen, China*

- **Generalizability of Multi-Agent RL**: To make RL algorithms generalizable in population-varying multi-agent systems, we proposed *Meta Representations for Agents* (MRA) that adopts multi-modal latent policies and a constrained mutual information maximization objective to discover the common strategic knowledge and diverse strategic modes. We proved that the learned policies can reach the Nash Equilibrium in every evaluation Markov game if with a sufficiently large latent space.
- **Visual Representation with Structured Data**: For structured visual tasks including person re-id and face recognition, we proposed to formulate feature interactions in a structured manner by taking advantage of the natural characteristics of data. Our *Structure-Regularized Attention* first captures informative patterns between neighbor nodes. Higher-level contextual information can then be accessed to enhance the desired features.

TEACHING EXPERIENCE

Head TA of CS 7648: Interactive Robot Learning (Fall 2021) at Georgia Tech.

SELECTED PROJECTS

Object Detection *May 2019 - Oct. 2019*
Project paper: Coarse-to-Fine Attention, advised by Bo Wu *Columbia University*

Cloth Simulation using OpenGL Shader *Jan. 2019 - May 2019*
Project website: ffjmmm.github.io/CS184-final/webpage, advised by Ren Ng *UC Berkeley*

PROFESSIONAL SERVICE

Conference Review: NeurIPS 2020/2021/2022, ICLR 2022/2023, AISTATS 2022/2023, RSS 2021, ICML 2022.
Journal Review: Neurocomputing.

HONORS AND AWARDS

NeurIPS Scholar Award *2022*
Georgia Tech Level A Premier Merit-Based Scholarship *2020-2021*
Second Prize in 2018 Undergraduate Electronics Design Contest *2018*
Third Prize in 2018 Intel Undergraduate Embedded System Contest *2018*
Outstanding Freshman Scholarship (Awarded to 30 among 6,500 students) *2016*