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 the Department of EECS, GPA: 3.90/4.00   *Berkeley, CA*

RESEARCH INTERESTS

My research centers around reinforcement learning (RL) and large language models (LLM). I'm interested in developing data-efficient RL algorithms that have both theoretical guarantees and strong empirical performance with application to robotic and multi-agent systems.

PUBLICATIONS AND PREPRINTS

[1] **Shenao Zhang**, Boyi Liu, Zhaoran Wang, Tuo Zhao, "Model-Based Reparameterization Policy Gradient: Theory and Practical Algorithms", under review. [PDF].

[2] **Shenao Zhang**, Wanxin Jin, Zhaoran Wang, "Adaptive Barrier Smoothing for First-Order Policy Gradient with Contact Dynamics", *Fortieth International Conference on Machine Learning (ICML)*, 2023. [PDF]

[3] **Shenao Zhang**, "Conservative Dual Policy Optimization for Efficient Model-Based Reinforcement Learning", *Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS)*, 2022. [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. [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** [3]: We studied the First-Order Poliy Gradient (FOPG) with complementarity-based models for robotic systems that experience hard contact. We showed that the convergence and gradient variance of FOPG are closely related to the model stiffness, which is determined by the centering parameter when solving the complementarity problem. Therefore, we proposed *Contact-Aware Analytic Barrier Smoothing*. By establishing the connection between analytic smoothing and randomized smoothing, we showed that the best linear approximation of the dynamics is captured by our method, whose gradient bias can thus be upper bounded.

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

- **Model-Based Reparameterization Policy Gradient Methods** [2]: We established the first convergence result for model-based RP PGMs, and our theory identified the smoothness of the function

approximators as a major determining factor that affects the quality of gradient estimation. Based on our theory, we proposed a spectral normalization method, which can effectively mitigate the exploding variance due to long model unrolls. Experimental results also support our theory and method.

- **Conservative Dual Policy Optimization for Efficient Model-Based RL** [1]: The provable efficiency results in previous Bayesian MBRL methods are statistically intractable in nonlinear settings since the assumption of restricted model complexity rarely holds. 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. Experiments were conducted in tabular and MuJoCo tasks.

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

- **Generalizability of Multi-Agent RL** [4]: 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 of Structured Data** [5]: For deformable objects, we proposed to take advantage of their natural characteristics by formulating the feature interactions in a structured manner. Our *Structure-Regularized Attention* captured informative patterns between neighborhood nodes. Higher-level contextual information can then be accessed to enhance the desired features.

South China University of Technology *May 2018 - Dec. 2018*
Undergraduate Researcher. Advisor: Huabiao Qin *Guangzhou, China*

- **Gaze Tracking Algorithms** [6]: We proposed gaze tracking algorithms based on projective mapping correction and gaze point compensation that work in the natural light environment.

TEACHING EXPERIENCE

Head TA of the graduate course 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. Related patent. *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 2020/21/22, ICLR 2022/23, AISTATS 2022/23, RSS 2021, ICML 2022/23.
Journal Review: Neurocomputing, Transactions on Pattern Analysis and Machine Intelligence (TPAMI).

HONORS AND AWARDS

NeurIPS Scholar Award	*2022*
Georgia Tech Level A Premier Merit-Based Scholarship	*2020-2021*
SCUT Study Abroad Global Education Scholarship	*2019*
Second Prize in the China Undergraduate Electronics Design Contest	*2018*
Third Prize in the Intel Undergraduate Embedded System Contest	*2018*
Outstanding Freshman Scholarship (Awarded to 30 among 6,500 students)	*2016*