Shangtong Zhang

422 Rice Hall, 85 Engineer's Way, Charlottesville, VA, 22903, United States
shangtong@virginia.edu https://shangtongzhang.github.io

Research Interest

The goal of my research is to solve sequential decision making problems in a scalable and reliable way. Currently, I focus on reinforcement learning as a solution method.

Academic Employments

Assistant Professor Aug. 2022 - Present
Department of Computer Science
University of Virginia, United States

Education

University of Oxford, United Kingdom Oct. 2018 - July. 2022
Doctor of Philosophy in Computer Science
Advisor: Prof. Shimon Whiteson

University of Alberta, Canada Sept. 2016 - Aug. 2018
Master of Science in Computer Science,
Advisor: Prof. Richard S. Sutton

Fudan University, China Sept. 2012 - Jun. 2016
Bachelor of Science in Computing Science

Publications

1. Direct Gradient Temporal Difference Learning
Xiaochi Qian[*advisee*], **Shangtong Zhang**.
arXiv:2308.01170, 2023.
2. Improving Monte Carlo Evaluation with Offline Data
Shuze Liu[*advisee*], **Shangtong Zhang**.
arXiv:2301.13734, 2023.
3. On the Convergence of SARSA with Linear Function Approximation
Shangtong Zhang, Remi Tachet des Combes, Romain Laroche.
International Conference on Machine Learning (**ICML**), 2023.
Acceptance rate: 28%
4. Global Optimality and Finite Sample Analysis of Softmax Off-Policy Actor Critic under State Distribution Mismatch
Shangtong Zhang, Remi Tachet des Combes[‡], Romain Laroche[‡].
Journal of Machine Learning Research (**JMLR**), 2022.
5. Truncated Emphatic Temporal Difference Methods for Prediction and Control
Shangtong Zhang, Shimon Whiteson.
Journal of Machine Learning Research (**JMLR**), 2022.
6. A Deeper Look at Discounting Mismatch in Actor-Critic Algorithms
Shangtong Zhang, Romain Laroche, Harm van Seijen, Shimon Whiteson, Remi Tachet des Combes.
International Conference on Autonomous Agents and Multiagent Systems (**AAMAS**), 2022.
Acceptance rate: 26%
7. Learning Expected Emphatic Traces for Deep RL
Ray Jiang, **Shangtong Zhang**, Veronica Chelu, Adam White, Hado van Hasselt.

AAAI Conference on Artificial Intelligence (**AAAI**), 2022.
Acceptance rate: 15%.

8. StarCraft II Unplugged: Large Scale Offline Reinforcement Learning
Michael Mathieu*, Sherjil Ozair*, Srivatsan Srinivasan*, Caglar Gulcehre*, **Shangtong Zhang***, Ray Jiang*, Tom Le Paine*, Richard Powell, Konrad Zolna, Julian Schrittwieser, David Choi, Petko Georgiev, Daniel Kenji Toyama, Aja Huang, Roman Ring, Igor Babuschkin, Timo Ewalds, Mahyar Bordbar, Sarah Henderson, Sergio Gomez Colmenarejo, Aaron van den Oord, Wojciech M. Czarnecki, Nando de Freitas, Oriol Vinyals.
arXiv:2308.03526, 2023
Deep RL Workshop at NeurIPS, 2021
9. Breaking the Deadly Triad with a Target Network
Shangtong Zhang, Hengshuai Yao, Shimon Whiteson.
International Conference on Machine Learning (**ICML**), 2021.
Acceptance rate: 21.5%.
10. Average-Reward Off-Policy Policy Evaluation with Function Approximation
Shangtong Zhang*, Yi Wan*, Richard S. Sutton, Shimon Whiteson.
International Conference on Machine Learning (**ICML**), 2021.
Acceptance rate: 21.5%.
11. Mean-Variance Policy Iteration for Risk-Averse Reinforcement Learning
Shangtong Zhang, Bo Liu, Shimon Whiteson.
AAAI Conference on Artificial Intelligence (**AAAI**), 2021.
Acceptance rate: 21.4%.
12. Learning Retrospective Knowledge with Reverse Reinforcement Learning
Shangtong Zhang, Vivek Veeriah, Shimon Whiteson.
Conference on Neural Information Processing Systems (**NeurIPS**), 2020.
Acceptance rate: 20.1%.
13. GradientDICE: Rethinking Generalized Offline Estimation of Stationary Values
Shangtong Zhang, Bo Liu, Shimon Whiteson.
International Conference on Machine Learning (**ICML**), 2020.
Acceptance rate: 21.8%.
14. Provably Convergent Two-Timescale Off-Policy Actor-Critic with Function Approximation
Shangtong Zhang, Bo Liu, Hengshuai Yao, Shimon Whiteson.
International Conference on Machine Learning (**ICML**), 2020.
Acceptance rate: 21.8%.
15. Deep Residual Reinforcement Learning
Shangtong Zhang, Wendelin Boehmer, Shimon Whiteson.
International Conference on Autonomous Agents and Multiagent Systems (**AAMAS**), 2020.
Acceptance rate: 23%.
Best Paper Award.
16. Mega-Reward: Achieving Human-Level Play without Extrinsic Rewards
Yuhang Song, Jianyi Wang, Thomas Lukasiewicz, Zhenghua Xu, **Shangtong Zhang**, Andrzej Wojcicki, Mai Xu
AAAI Conference on Artificial Intelligence (**AAAI**), 2020.
Acceptance rate: 20.6%.
17. DAC: The Double Actor-Critic Architecture for Learning Options
Shangtong Zhang, Shimon Whiteson.
Conference on Neural Information Processing Systems (**NeurIPS**), 2019.
Acceptance rate: 21.2%.

18. Generalized Off-Policy Actor-Critic
Shangtong Zhang, Wendelin Boehmer, Shimon Whiteson.
Conference on Neural Information Processing Systems (**NeurIPS**), 2019.
Acceptance rate: 21.2%.

19. Distributional Reinforcement Learning for Efficient Exploration
Borislav Mavrin, **Shangtong Zhang**[†], Hengshuai Yao, Linglong Kong, Kaiwen Wu, Yaoliang Yu
International Conference on Machine Learning (**ICML**), 2019.
Acceptance rate: 22.6%.
A short version is accepted as an extended abstract at AAMAS 2019.

20. ACE: An Actor Ensemble Algorithm for Continuous Control with Tree Search
Shangtong Zhang, Hao Chen, Hengshuai Yao.
AAAI Conference on Artificial Intelligence (**AAAI**), 2019.
Acceptance rate: 16.2%.

21. QUOTA: The Quantile Option Architecture for Reinforcement Learning
Shangtong Zhang, Borislav Mavrin, Linglong Kong, Bo Liu, Hengshuai Yao.
AAAI Conference on Artificial Intelligence (**AAAI**), 2019.
Acceptance rate: 16.2%.

22. MLPack 3: A Fast, Flexible Machine Learning Library
Ryan Curtin, Marcus Edel, Mikhail Lozhnikov, Yannis Mentekidis, Sumedh Ghaisas, **Shangtong Zhang**
Journal of Open Source Software (**JOSS**), 2018.

23. Crossprop: Learning Representations by Stochastic Meta-Gradient Descent in Neural Networks
Vivek Veeriah[*], **Shangtong Zhang**[*], Richard S. Sutton.
European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML-PKDD**), 2017.
Acceptance rate: 27.1%.

24. A Deeper Look at Experience Replay
Shangtong Zhang, Richard S. Sutton.
Deep RL Symposium at NIPS, 2017.

25. Comparing Deep Reinforcement Learning and Evolutionary Methods in Continuous Control
Shangtong Zhang, Osmar R. Zaiane
Deep RL Symposium at NIPS, 2017.

26. A Demon Control Architecture with Off-Policy Learning and Flexible Behavior Policy
Shangtong Zhang, Richard S. Sutton.
Hierarchical RL Workshop at NIPS, 2017.

27. A Deep Neural Network for Modeling Music
Pengjing Zhang, Xiaoqing Zheng, Wenqiang Zhang, Siyan Li, Sheng Qian, Wenqi He, **Shangtong Zhang**, Ziyuan Wang
International Conference on Multimedia Retrieval (**ICMR**), 2015.
Acceptance rate: 31%.

[*]: Equal contribution
[‡]: Equal advising
[†] : My name does not appear in the ICML proceedings due to a mistake in submission. See Acknowledgments, arXiv, or AAMAS proceedings for clarification.

Services

Meta Reviewer & Area Chair
ICLR 2024
AISTATS 2024
ACML 2022, 2023

Reviewer & Program Committee
Transactions on Pattern Analysis and Machine Intelligence (1)
Transaction of Machine Learning Research (2)
Journal of Machine Learning Research (3)
Artificial Intelligence Journal (2)
Transactions on Intelligent Systems and Technology (2)
IJCAI 2023
AISTATS 2022
NeurIPS 2020, 2021, 2022, 2023
ICML 2020, 2021, 2022, 2023
AAAI 2020, 2021, 2022, 2023
ICLR 2021, 2022, 2023
SIGCOMM 2022
Offline Reinforcement Learning Workshop at NeurIPS 2020, 2021, 2022
Deep Reinforcement Learning Workshop at NeurIPS 2019, 2020, 2021, 2022
Adaptive and Learning Agents Workshop at AAMAS 2019, 2020
Optimization Foundations for Reinforcement Learning Workshop at NeurIPS 2019
Reinforcement Learning for Real Life Workshop at ICML 2019, 2021
Reinforcement Learning for Real Life Workshop at NeurIPS 2022

Conference Session Chair
AAAI 2023, "Reinforcement Learning Theory & Algorithms"

Departmental Services
Graduate Admission Committee at CS of UVA, 2022 - 2023 AY

Honours

Runner-Up for the IFAAMAS Victor Lesser Dissertation Award	2022
Alf Weaver Junior Faculty Fellowship, University of Virginia	2022 - 2027
EPSRC Studentship, University of Oxford	2018 - 2022
AAMAS Student Scholarship	2022
ICLR Outstanding Reviewer	2021
NeurIPS Reviewer Award	2020
ICML Reviewer Award	2020
Light Senior Scholarship, St Catherine's College, University of Oxford	2020
AAMAS Travel Award	2020
AAMAS Best Paper Award	2020
NeurIPS Optimization Foundations for RL Workshop Travel Award	2019
NeurIPS Travel Award	2019
AAAI Travel Award	2019
NIPS Hierarchical RL Workshop Travel Award	2017
EMC Scholarship, Fudan University	2014

Supervision

PhD Students
Ethan Blaser (2023 -)
Jiuqi Wang (2023 -)
Shuze Liu (2022 -)

Master Students
Licheng Luo (2023 -)
Kefan Song (2023 -)

Undergraduates
Ja-Zhua Cheng (2022 -)

Research Assistants
Xiaochi Qian (2022 -)

Alumni
Zhengkun Xiao (Master, 2022 - 2023, now PhD student at University of Florida)
Jiuqi Wang (RA, 2022 - 2023, now PhD student at University of Virginia)

PhD Committees
Ingy ElSayed-Aly (by Prof. Lu Feng at UVA)
Sudhir Shenoy (by Prof. Afsaneh Doryab at UVA)
Chuanhao Li (by Prof. Hongning Wang at UVA)
Qualification:
Zeyu Mu (by Prof. Brian Park at UVA)
Dane Williamson (by Prof. Yangfeng Ji at UVA)
Matthew Landers (by Prof. Afsaneh Doryab at UVA)
Ethan Harrison Blaser (by Prof. Hongning Wang at UVA)

Invited Talks

Offline Reinforcement Learning: Current and Future 2023
AAAI New Faculty Highlight Program

Breaking the Deadly Triad in Off-Policy Reinforcement Learning
School of Computing Science, Simon Fraser University 2022
Department of Electrical & Computer Engineering, University of Waterloo 2022
Department of Computer Science, University of Virginia 2022
School of Informatics, University of Edinburgh 2021

Breaking the Deadly Triad in Reinforcement Learning 2021
RL team, DeepMind, hosted by Hado van Hasselt

Breaking the Deadly Triad with a Target Network 2021
Microsoft Research Summit

Off-Policy Evaluation 2020
Data Fest 2020, Open Data Science

Off-Policy Evaluation and Control 2020
ByteDance AI Lab, Shanghai

Off-Policy Actor-Critic Algorithms 2019
Latent Logic LTD, Oxford

Generalized Off-Policy Actor-Critic 2019
Noah's Ark Lab, Huawei, Edmonton

Exploration with Quantile Options 2018
Huawei RL Workshop, Edmonton

Coding Deep RL Papers 2017
NIPS MLTrain Workshop, Long Beach

Teaching

University of Virginia, Instructor
CS6501: Topics in Reinforcement Learning Fall 2022

University of Oxford, Teaching Assistant Michaelmas 2019
AIMS CDT Lectures

University of Alberta, Teaching Assistant Fall 2016
CMPUT 229 Computer Organization and Architecture

Research Internships

Microsoft Research Montreal, Canada Jun. 2021 - Sept. 2021
Collaboration: Remi Tachet des Combes, Romain Laroche, and Harm van Seijen

DeepMind London, United Kingdom Feb. 2021 - Jun. 2021
Collaboration: AlphaStar team (Michael Mathieu, Oriol Vinyals, etc)
Collaboration: Adam White and Hado van Hasselt

DeepDrive, Edmonton, Canada Sept. 2020 - Dec. 2020
Collaboration: Hengshuai Yao

Microsoft Research Montreal, Canada Jun. 2020 - Aug. 2020
Collaboration: Remi Tachet des Combes, Romain Laroche, and Harm van Seijen

Noah's Ark Lab, Huawei, Edmonton, Canada May. 2018 - Aug. 2018
Collaboration: Hengshuai Yao

Code

PyTorch Deep RL
A zoo of popular deep RL algorithms in PyTorch with **2.8k stars** in Github.

Reinforcement Learning: An Introduction
Python implementation of the book *Reinforcement Learning: An Introduction* with **11.7k stars** in Github.

Google Summer of Code (GSoC) 2017
Contributed to MLPack by implementing a deep RL framework.

Google Summer of Code (GSoC) 2014
Contributed to Xapian by optimizing the post list and the position list.