# Ishan Durugkar


## Senior Research Scientist, Sony AI • Reinforcement Learning

@ ishand@cs.utexas.edu  Greater Seattle, WA  idurugkar.github.io  Google Scholar  github.com/idurugkar
in/idurugkar


### Summary


Reinforcement learning researcher at Sony AI who both **ships and publishes**: I train, polish, and deploy RL agents inside large AAA video games under real-time, human-in-the-loop constraints, and I publish foundational RL research at NeurIPS, ICLR, ICML, and AAAI. I completed my Ph.D. at the University of Texas at Austin, advised by Prof. Peter Stone, with a thesis on improving RL algorithms by controlling the data they generate. My interests span deep, multi-agent, imitation, goal-conditioned, and unsupervised RL, and increasingly how LLMs and agentic tooling can make learned policies more capable and accessible.


### Selected Highlights

- **Deployed RL in shipped-scale games.** Train RL agents for complex tasks in (undisclosed) AAA Sony titles — on the same Sony AI Game AI team behind *GT Sophy*, the Gran Turismo champion-level agent.
- **Real-world deployment under hard constraints.** Polish and deploy research agents for internal demos where policies must run in real time alongside a live game, handle noisy real-time human interaction, and where I evaluate and select the best policy snapshots to present.
- **LLMs × RL.** Built MCP tools and an LLM interface to interact with and control trained RL policies in-game, making learned policies far more accessible.
- **Foundational research, broadly cited.** *Generative Multi-Adversarial Networks* (ICLR 2017), an early and widely-cited multi-adversary GAN framework; first-author work at NeurIPS, ICLR, and AAAI.

### Research & Work Experience

**Research Scientist**
**Sony AI — Game AI** 2023 +
On the Game AI team conducting fundamental RL research and applying it to create more capable AI agents for video games. I extract features from games, design rewards, and tune algorithms to train RL agents that accomplish interesting and complex tasks in (undisclosed) AAA titles, and contribute to research efforts that push the envelope of what is possible in these massive games. I also own the path from research to demo — deploying policies that must run in real time alongside a live game, remain robust to noisy real-time interaction with people, and selecting the best policy snapshots to present. Most recently I built MCP tools and an LLM interface to interact with and control trained RL policies in-game, making learned policies more accessible. (Same team behind *GT Sophy*, the Gran Turismo champion-level agent.)

**Research Scientist Intern**
**DeepMind** Summer 2021
Mentored by Volodymyr Mnih in the Discovery team. Research and development of deep RL algorithms, specifically unsupervised RL and learning of skillful policies in the absence of reward signals from the environment. Evaluation of the learned skills on the Atari benchmark.

**Research Intern**
**Microsoft Research (MSR)** Summer 2018
With Adith Swaminathan and Matthew Hausknecht, research into approaches to improve the ease of developing RL agents. Focused on incorporating multiple feedback channels into learning the agent policy; specifically, leveraged demonstrations to speed up learning of a policy-gradient-based agent.

**Graduate Research Assistant**
**The University of Texas at Austin** 2017–2023
Ph.D. research advised by Peter Stone (LARG). Primary focus on Reinforcement Learning, with projects on optimization of TD learning, human feedback, sim-to-real transfer, multi-agent RL, and goal-conditioned RL. Doctoral thesis: improving RL algorithms by estimating and controlling the visitation distributions induced by the agent.

**Research Assistant**
**University of Massachusetts at Amherst** 2015–2017
Research with Sridhar Mahadevan (ALL) on: *(1)* Generative Adversarial Networks (GANs) and Variational Auto-Encoders (VAEs), *(2)* temporal abstractions for actions in Deep RL, *(3)* manifold learning on graphs, and *(4)* learning policies to navigate large knowledge bases.

**Research Member**
**Maharashtra Institute of Technology** 2013
Research on non-convex optimization using swarm techniques.

### Selected Publications

(**bold** = author; see the **Publications** section below for the full list)

- Jacob Berg, Chuning Zhu, Yanda Bao, **Ishan Durugkar**, Abhishek Gupta. "Semantic World Models." *arXiv:2510.19818*, 2025.
- Caroline Wang, Muhammad A. Rahman, **Ishan Durugkar**, Elad Liebman, Peter Stone. "N-Agent Ad Hoc Teamwork." **NeurIPS** 2024.
- Siddhant Agarwal, **Ishan Durugkar**, Peter Stone, Amy Zhang. "f-Policy Gradients: A General Framework for Goal-Conditioned RL using f-Divergences." **NeurIPS** 2023.
- Caroline Wang, **Ishan Durugkar**, Elad Liebman, Peter Stone. "$DM^2$: Distributed Multi-Agent Reinforcement Learning via Distribution Matching." **AAAI** 2023.
- **Ishan Durugkar**, Mauricio Tec, Scott Niekum, Peter Stone. "Adversarial Intrinsic Motivation for Reinforcement Learning." **NeurIPS** 2021.
- Siddharth Desai, **Ishan Durugkar**, Haresh Karnan, Garrett Warnell, Josiah Hanna, Peter Stone. "An Imitation from Observation Approach to Transfer Learning with Dynamics Mismatch." **NeurIPS** 2020. *(joint first author)*
- Brahma Pavse, **Ishan Durugkar**, Josiah Hanna, Peter Stone. "Reducing Sampling Error in Batch Temporal Difference Learning." **ICML** 2020.
- Rajarshi Das, Shehzaad Dhuliawala, Manzil Zaheer, Luke Vilnis, **Ishan Durugkar**, Akshay Krishnamurthy, Alex Smola, Andrew McCallum. "Go for a Walk and Arrive at the Answer: Reasoning over Paths in Knowledge Bases using RL." **ICLR** 2018.
- **Ishan Durugkar**, Ian Gemp, Sridhar Mahadevan. "Generative Multi-Adversarial Networks." **ICLR** 2017.

## Education

### Ph.D. Computer Science (2017-2023)
**The University of Texas at Austin, USA**
Research areas: Reinforcement Learning, Multi-agent Reinforcement Learning, Robotics, Imitation Learning, Deep Learning
Affiliations: Peter Stone, Learning Agents Research Group (LARG)

### M.S. Computer Science (2014 – 2017)
**University of Massachusetts Amherst**
Research Areas: Deep Learning, Reinforcement Learning, Generative Adversarial Networks
GPA: 3.97
Affiliations: Sridhar Mahadevan, Autonomous Learning Lab (ALL)

### B.Engg. Computer Engineering (2007 – 2013)
**Maharashtra Institute of Technology, India**
Research and Extracurricular interests: Robotics, Swarm techniques, non-convex optimization

## Skills

**Reinforcement learning:** deep RL, multi-agent RL, imitation & inverse RL, goal-conditioned RL, unsupervised RL / skill discovery, sim-to-real transfer; reward design and RL-policy deployment.
**LLMs & agents:** LLM-driven control of RL policies, MCP tool building, agentic interfaces.
**Engineering:** Python (proficient); C++ (intermediate); PyTorch, JAX, TensorFlow (and Theano); large-scale / distributed NN training (Slurm, HTCondor, and proprietary systems); HPC clusters.

## Research Statement

I am a Research Scientist at Sony AI working in the Game AI team. We focus on fundamental **Reinforcement Learning** (RL) research and apply techniques to create more capable AI agents for video games. Challenges we focus on include RL training techniques for agents that can be deployed on consumer PlayStation devices, are predictable and enjoyable to interact with, and empower developers to create the next generation of video games.

I also contribute to research threads focusing on **multi-agent reinforcement learning**, **robotics**, and **goal-conditioned reinforcement learning**, with collaborators at UT Austin and the University of Washington. I completed my Ph.D. from the Department of Computer Science at the University of Texas at Austin, advised by Prof. Peter Stone. My doctoral thesis focused on improving RL algorithms by controlling the data they generate — in particular, the estimation and control of the visitation distributions induced by RL agents, applied to imitation learning, goal-conditioned RL, sim-to-real transfer, and multi-agent RL.

My previous industry experience has been in research internships at **DeepMind** in 2021 and **Microsoft Research** in 2018.

## Other Work Experience

**Software Engineering Intern**
**Amazon Comixology, New York City** Summer 2015
Implemented components of the Comixology continuous integration pipeline. Developed a build system based on Ruby, PHP and ANT that enabled developers to run Unit, Integration and Functional Tests in a single setting, and integrated it into the development process.

**Data Scientist – Research Assistant**
**UMass Amherst Alumni Association** 2015 – 2016
Used data aggregation, cleaning and analysis along with machine learning techniques to identify and better engage alumni with on-campus events and fund raising.

**Software Engineer**
**Microsoft India, Hyderabad** 2013 – 2014
As part of the Code Search team in Visual Studio Online, implemented a prototype of code search based on Elasticsearch, which was then approved for further development. Participated in the effort to scale the prototype to integrate with Visual Studio Online.

## Robotics Competitions

**RoboCup Standard Platform League**
**UT Austin Villa** 2019 - 2022
Quarter-finalists 2019, 2022, 4th place in 2021
**ABU Asia-Pacific Robot Contest**
**Maharashtra Institute of Technology** 2010 – 2012
Represented India in 2010 and 2012

## Teaching

**Teaching Assistant**
**The University of Texas at Austin** Fall 2019
Reinforcement Learning: Theory and Practice
**Teaching Assistant**
**UMass Amherst** 2015 – 2017
- Machine Learning
- Distributed and Operating Systems

## Service

**Area Chair**
**ICML, RLC** 2025

**Organizer – GCRL Workshop**
**NeurIPS** Fall 2024
Member of the organizing committee for the goal-conditioned RL workshop at NeurIPS 2023 `https://goal-conditioned-rl.github.io/2023/`

**RoboCup Standard Platform League Organizing Committee**
**UT Austin** Fall 2022 - Spring 2023
Member of the organizing committee for the standard platform league (SPL) at RoboCup 2023 to be held in Bordeaux, France `https://spl.robocup.org/committees-2023/`

**International Joint Conference on Artificial Intelligence (IJCAI)**
**Assistant Workflow Chair** 2023

**RL Reading Group – Organizer**
**UT Austin** Spring 2019 – Fall 2022

**Graduate Admission Committee**
**UT Austin** Spring 2019
Volunteer

**Indian Student Association**
**University of Massachusetts, Amherst** Fall 2015 – Fall 2016
Vice-President of the Indian Student Association at the University of Massachusetts, Amherst.

## Reviewing Experience

**Conferences**: AAAI (2021–2023), ICLR (2021–2025, *Highlighted Reviewer* 2022), NeurIPS (2016, 2020–2024), ICML (2021, 2022, 2024, 2025), CoLLAs (2022, *inaugural PC*), AISTATS (2022, 2023), CoRL (2020), IJCAI (2020, 2021), ICRA (2023), RLC (2024 *inaugural PC*, 2025).

**Journals**: JAIR (2022), JMLR (2023, 2024), TMLR (2022–2024), IEEE TNNLS (2020–2023), Springer Machine Learning (2022–2024).

**Workshops**: Lifelong Machine Learning (ICML 2020), Multi-Task and Lifelong RL (ICML 2019), Neverending Learning (ICLR 2021).

## Publications

### Conference Proceedings


[Wan+24] Caroline Wang et al. "N-agent ad hoc teamwork". In: *Advances in Neural Information Processing Systems*. Vol. 37. 2024, pp. 111832–111862.

[Aga+23] Siddhant Agarwal et al. "f-Policy Gradients: A General Framework for Goal-Conditioned RL using f-Divergences". In: *Advances in Neural Information Processing Systems*. Vol. 36. 2023.

[Wan+23] Caroline Wang et al. "DM $^2$: Distributed Multi-Agent Reinforcement Learning via Distribution Matching". In: *Proceedings of the $37^{th}$ AAAI Conference on Artificial Intelligence*. 2023.

[Nar+22] Sai Kiran Narayanaswami et al. "Towards a real-time, low-resource, end-to-end object detection pipeline for robot soccer". In: *Robot World Cup*. **nominated for best paper**. Springer, 2022, pp. 62–74.

[Dur+21a] Ishan Durugkar et al. "Adversarial Intrinsic Motivation for Reinforcement Learning". In: *Advances in Neural Information Processing Systems*. 2021.

[Des+20] Siddharth Desai et al. "An Imitation from Observation Approach to Transfer Learning with Dynamics Mismatch". In: *Advances in Neural Information Processing Systems*. *Joint first author*. 2020.

[DLS20] Ishan Durugkar, Elad Liebman, and Peter Stone. "Balancing Individual Preferences and Shared Objectives in Multiagent Reinforcement Learning". In: *Proceedings of the $29^{th}$ International Joint Conference on Artificial Intelligence*. 2020, pp. 2505–2511.

[Pav+20b] Brahma Pavse et al. "Reducing Sampling Error in Batch Temporal Difference Learning". In: *Proceedings of $37^{th}$ International Conference on Machine Learning, Vienna, Austria*. 2020.

[Das+18] Rajarshi Das et al. "Go for a Walk and Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning". In: *International Conference on Learning Representations*. 2018.

[DGM17] Ishan Durugkar, Ian Gemp, and Sridhar Mahadevan. "Generative multi-adversarial networks". In: *International Conference on Learning Representations*. 2017.

[DS17a] Ishan Durugkar and Peter Stone. "TD Learning with Constrained Gradients". In: *Proceedings of the Deep Reinforcement Learning Symposium, NIPS 2017*. Long Beach, CA, USA, Dec. 2017.

[PGD17] Mario Parente, Ian Gemp, and Ishan Durugkar. "Unmixing in the presence of nuisances with deep generative models". In: *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE. 2017, pp. 5189–5192.

[Tho+17] Philip S Thomas et al. "Predictive off-policy policy evaluation for nonstationary decision problems, with applications to digital marketing". In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. 2017, pp. 4740–4745.

[KDK13] Anand J Kulkarni, Ishan P Durugkar, and Mrinal Kumar. "Cohort intelligence: a self supervised learning behavior". In: *2013 IEEE international conference on systems, man, and cybernetics*. IEEE. 2013, pp. 1396–1400.


### Workshops and Symposium


[Hud+22] Eddy Hudson et al. "ABC: Adversarial Behavioral Cloning for Offline Mode-Seeking Imitation Learning". arXiv preprint arXiv:2211.04005, Deep RL Workshop and Offline RL workshop at NeurIPS. 2022.

[Dur+21b] Ishan Durugkar et al. "Wasserstein Distance Maximizing Intrinsic Control". NeurIPS Deep Reinforcement Learning Workshop. 2021.

[Pav+20a] Brahma S. Pavse et al. "On Sampling Error in Batch Action-Value Prediction Algorithms". In the Offline Reinforcement Learning Workshop at Neural Information Processing Systems (NeurIPS), December 2020. Dec. 2020.

[Dur+19] Ishan Durugkar et al. "Multi-Preference Actor Critic". The Multi-disciplinary Conference on Reinforcement Learning and Decision Making. 2019.

[DS18] Ishan Durugkar and Peter Stone. "Adversarial goal generation for intrinsic motivation". AAAI Student Abstracts. 2018.

[DS17b] Ishan Durugkar and Peter Stone. "TD learning with constrained gradients". NeurIPS Deep Reinforcement Learning Symposium. 2017.


### Preprints


[Ber+25] Jacob Berg et al. *Semantic World Models*. 2025.

[Wan+25] Caroline Wang et al. *Sequence Modeling for N-Agent Ad Hoc Teamwork*. 2025.

[Dur+16] Ishan P Durugkar et al. *Deep reinforcement learning with macro-actions*. arXiv preprint 1606.04615. 2016.

[Gem+16a] Ian Gemp et al. *Deep Generative Models for Spectroscopic Analysis on Mars*. ResearchGate. 2016.

[Gem+16b] Ian Gemp et al. *Inverting variational autoencoders for improved generative accuracy*. arXiv preprint 1608.05983. 2016.