

I conduct research in reinforcement learning, with a focus on temporal abstraction and spatial abstraction.

Education

- 2021 - Present **Ph.D. in Computer Science**, *University of Massachusetts, Amherst*, MA, USA.
Supervisors: Bruno Castro Da Silva & Philip Thomas, Autonomous Learning Lab (ALL)
GPA: 4.0/4.0
- 2019 - 2021 **M.Sc. in Computing Science**, *University of Alberta*, AB, Canada.
Supervisor: Martha White, RLAI, AMII
GPA: 4.0/4.0
- 2015-2019 **B.Tech in Computer Science and Engineering**, *IIT*, Patna, India.
Supervisors: Sriparna Saha & Pushpak Bhattacharyya, AI-ML-NLP Lab
CPI: 9.49/10 , Department Rank: 3/55

Publications

- [C7] **Gupta D.***, Chandak, Y.*, Jordan, S. M., Thomas, P. S., & Silva, B.C. da. Behavior Alignment via Reward Function Optimization. (**Spotlight**) *To Appear in Neural Information Processing Systems (NeurIPS)* 2023.
- [C6] **Gupta D.**, Chow, Y., Tuleberggenov, A., Ghavamzadeh, M. & Boutilier, C. Offline Reinforcement Learning for Mixture-of-Expert Dialogue Management. *To Appear in Neural Information Processing Systems (NeurIPS)* 2023.
- [C5, W3] Chow, Y., Tuleberggenov, A., Nachum, O., **Gupta D.**, Ryu, M., Ghavamzadeh, M. & Boutilier, C. A Mixture-of-Expert Approach to RL-based Dialogue Management, *International Conference on Learning Representations (ICLR)*, 2023 & *Workshop on Foundation Models for Decision Making (NeurIPS)*, 2022.
- [W2] Kostas, J., Jordan, S., Chandak, Y., Theocharous, G., **Gupta, D.**, & Thomas, P. A Generalized Learning Rule for Asynchronous Coagent Networks, *Reinforcement Learning and Decision Making (RLDM)*, 2022.
- [C4] **Gupta, D.**, Mihucz, G., Schlegel, M., Kostas, J., Thomas, P., & White, M. Structural Credit Assignment in Neural Networks using Reinforcement Learning. *Neural Information Processing Systems (NeurIPS)*, 2021.
- [J4] Saha, T.*, **Gupta, D.***, Saha, S., & Bhattacharyya, P. A Unified Dialogue Management Strategy for Multi-intent Dialogue Conversations in Multiple Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 2021.
- [W1] **Gupta, D.**, Schlegel, M., Kostas, J., Mihucz, G., & White, M. Investigating Coagent Networks for Supervised Learning, *BeyondBackprop Workshop, Neural Information Processing Systems(NeurIPS)*, 2020.
- [C3] Ghiassian, S.*, Patterson, A.*, Garg, S., **Gupta, D.**, White, A., & White, M. Gradient Temporal-Difference Learning with Regularized Corrections. *International Conference on Machine Learning (ICML)* 2020.
- [J3] Saha, T.*, **Gupta, D.***, Saha, S., & Bhattacharyya, P. Towards Integrated Dialogue Policy Learning for Multiple Domains and Intents using Hierarchical Deep Reinforcement Learning. *Expert Systems with Applications*, 2020.
- [J2] Saha, T., **Gupta, D.**, Saha, S., & Bhattacharyya, P. A hierarchical approach for efficient multi-intent dialogue policy learning. *Multimedia Tools and Applications*, 2020.
- [J1] Saha, T., **Gupta, D.**, Saha, S., & Bhattacharyya, P. Emotion Aided Dialogue Act Classification for Task-Independent Conversations in a Multi-modal Framework. *Cognitive Computation*, 2020.
- [C2] Saha, T., **Gupta, D.**, Saha, S., & Bhattacharyya, P. Reinforcement Learning Based Dialogue Management Strategy. In *L. Cheng, A. C. S. Leung, & S. Ozawa (Eds.), Neural Information Processing (ICONIP)*, 2018.

[C1] Agrawal, K., Jain, K., **Gupta, D.**, Srivastav, R., Agnihotri, A., & Thakur, A. Bayesian Optimization Based Terrestrial Gait Tuning for a 12-DOF Alligator-Inspired Robot With Active Body Undulation. ASME-IDETC, 2018.

Pre-Prints

[PP2] Sun, S., **Gupta, D.**, & Iyyer, M. Exploring the impact of low-rank adaptation on the performance, efficiency, and regularization of RLHF. *ArXiv:2309.09055*.

[PP1] Kostas, J. E., Jordan, S. M., Chandak, Y., Theocharous, G., **Gupta, D.**, White, M., Silva, B. C. da, & Thomas, P. S. Coagent Networks: Generalized and Scaled. *ArXiv:2305.09838*.

Patents

Easa, Z., **Gupta, D.**, Mathew, J. & Mathew, A. (2017). System and method for detecting a change in occupancy status of a slot over a platform. *Pending at Indian Patent Office: Application No. 201731036379*

Research Experience

- Jul 2023 – **Student Researcher, Google Research**, hosted by Dr. Yinlam Chow, Mountain View, CA.
Developing reinforcement learning techniques to enhance decoding in large language models (LLMs).
- Sep 2021 – **Research Assistant, UMass**, supervised by Dr. Bruno C.S. Silva & Dr. Philip Thomas, MA.
Working on reward alignment, credit assignment, and temporal abstraction problems. Work accepted at NeurIPS 2023.
- Jun 2022 – **Student Researcher, Google Research**, hosted by Dr. Yinlam Chow, Mountain View, USA.
Feb 2023 Worked on steerable language models using RL. Work published at ICLR 2023 and NeurIPS 2023.
- May 2020 – **Research Assistant, RLAI & AMII**, supervised by Dr. Martha White, Edmonton, Canada.
Aug 2021 Worked on off-policy learning and online learning architectures in RL. Work published at ICML 2020 and NeurIPS 2021.
- May 2018 – **Research Intern, IBM Research Labs**, hosted by Dr. Kedar Kulkarni, Bangalore, India.
Aug 2018 Worked on clustering time series data for premature anomaly detection.
- May 2017 – **Research Intern, IIT Delhi**, supervised by Dr. Sanjit Kaul & Dr. Saket Anand, Delhi, India.
July 2017 Used IRL to extract reward functions later applied for obstacle avoidance in self-driving cars.

Teaching Experience

- Sept 2022 – **Teaching Assistant, University of Massachusetts**, Amherst, USA.
Dec 2022 ◦ COMPSCI 687: Reinforcement Learning (Fall 22)
- Sept 2021 – **Teaching Assistant, University of Massachusetts**, Amherst, USA.
Dec 2021 ◦ COMPSCI 383: Artificial Intelligence (Fall 21)
- August 2019 – **Teaching Assistant, University of Alberta**, Edmonton, Canada.
May 2021 ◦ CMPUT 397: Introduction to RL (Fall 20)
◦ CMPUT 175: Introduction to Python - II (Fall 19 & Winter 20)

Technical skills

- **Programming Languages:** Python, C/C++, Bash, Embedded C, Latex
- **Tools:** JAX, PyTorch, Tensorflow, ROS, Matplotlib Numpy, Scikit-Learn, OpenCV, Gazebo.
- **Platforms:** MAC OS, Linux, Windows

Service and Extra-Curricular

- Reviewer: **NeurIPS 2023, ICML 2023, ICLR 2021, ICML 2020, TPAMI**.
- ML Faculty hiring interviews for UMass Amherst (Spring 2023).
- Helping organize a tutorial on policy optimization in reinforcement learning, **NeurIPS 2020**.
- Organizer, RL Social at **ICML 2020 & ICLR 2020**.
- Founder & Coordinator, Tinkerer's Lab, IIT Patna.
- Technical Secretary, Gymkhana, IIT Patna.
- Mentor, Robotics Lab, BCE Patna.