Brahma S. Pavse

Contact Information

Email: pavse@wisc.edu
Webpage: `https://brahmasp.github.io`
Google Scholar: `https://scholar.google.com/citations?user=2Dc_GnUAAAAJ&hl`

Education

University of Wisconsin - Madison (01/2022 - present)

- Ph.D. in Computer Science.
- Interests: reinforcement learning, representation learning, hierarchical RL.
- Advisor: Josiah P. Hanna.
- Cisco Systems Distinguished Graduate Fellow.

University of Texas at Austin (2015 - 2020)

- B.S and M.S. in Computer Science.
- Advisor: Peter Stone.

Academic Research Experience

University of Wisconsin – Madison, Madison, WI, USA
Research Assistant and Cisco Systems Distinguished Fellow **Spring 2022 -**

- Currently investigating how to learn effective state representations for planning in hierarchical reinforcement learning agents that learn from raw sensorimotor experience.
- Developed novel algorithms to improve the quality of representations for data-efficient and accurate evaluation of untested reinforcement learning agents.
- Developed techniques to improve robustness of reinforcement learning agents in queuing network problems with exogenous stochasticity and unbounded state spaces.
- Mentored three undergraduate students.

University of Texas at Austin, Austin, TX, USA
Research Assistant **August 2018 - August 2020**

- Developed algorithms to accurately and data-efficiently evaluate untested reinforcement learning agents.
- Developed an imitation from observation and reinforcement learning algorithm to speed up sample efficiency on a real robot.

Industry Research Experience

Netflix Research, Los Gatos, CA, USA
Machine Learning Research Intern — Recommendations team **Summer 2025**

- Mentor: Qitong Gao. Collaborator: Adith Swaminathan.
- Developed the first RL-based system for personalizing images of titles at Netflix.
- Implemented the data and training pipeline to train and evaluate offline RL algorithms such as SARSA and conservative Q-learning.
- Achieved comparable performance to a decade-old bandit system on short-term metrics and captured long-term return better than the bandit model by 33%.
- Currently undergoing A/B test evaluations.

Sony AI America, Remote, USA
AI Research Intern — Reinforcement learning team **Summer 2023**

- Mentor: Varun Kompella.

- Leveraged value iteration networks to improve the quality of representations learned to speedup sample efficiency of an RL agent.
- Simulation results showed improvements in data-efficiency over baseline agent.

Industry Engineering Experience

Salesforce.com, San Francisco, CA, USA
Software Engineer — Database Optimization team **Aug. 2020 - Jan. 2022**

Salesforce.com, San Francisco, CA, USA
Software Engineering Intern — Database Optimization team **Summer 2019, 2018, 2017**

SAS Institute, Cary, NC, USA
Software Engineering Intern — Data Management team **Summer 2016**

Awards and Honors

- **Cisco Systems Distinguished Graduate Fellowship (2025).**
- NeurIPS Top Reviewer Award (Top 8%) (2025, 2024, 2023).
- AAAI Student Scholarship (2023).
- **UW Madison CS Summer Research Fellowship Award (2022).**
- **UW Madison CS Graduate Scholarship (2022).**
- UT Austin CS Special Departmental Honors (Research) (2020).
- **Eva Stevenson Woods Endowed Presidential Scholarship (2019).**
- **National Instruments Endowed Scholarship (2019).**
- RoboCup 3D Simulation League World Champions (2019, 2018).

Publications (*= contribution)

Peer-reviewed Conference Papers

6. **Brahma S. Pavse**, Yudong Chen, Qiaomin Xie, and Josiah P. Hanna. Stable Offline Value Function Learning with Bisimulation-based Representations. International Conference on Machine Learning (ICML), July 2025. Acceptance rate: 26.9%. [PDF].
5. **Brahma S. Pavse**, Matthew Zurek, Yudong Chen, Qiaomin Xie, and Josiah P. Hanna. Learning to Stabilize Online Reinforcement Learning in Unbounded State Spaces. International Conference on Machine Learning (ICML), July 2024. Acceptance rate: 27.5%. [PDF].
4. **Brahma S. Pavse** and Josiah P. Hanna. State-Action Similarity-Based Representations for Off-Policy Evaluation. Neural Information Processing Systems (NeurIPS), December 2023. Acceptance rate: 26.1%. [PDF].
3. **Brahma S. Pavse** and Josiah P. Hanna. Scaling Marginalized Importance Sampling to High-Dimensional State-Spaces via State Abstraction. Association for the Advancement of Artificial Intelligence (AAAI), February 2023. Acceptance rate: 19.6%. **Selected for oral presentation**. [PDF].
2. **Brahma S. Pavse***, Faraz Torabi*, Josiah P. Hanna, Garrett Warnell, Peter Stone. RIDM: Reinforced Inverse Dynamics Modeling for Learning From a Single Observed Demonstration. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 2020. Acceptance rate: 47%. **2nd place in the RoboCup 3D Sim Scientific Challenge 2019**. [PDF].
1. **Brahma S. Pavse**, Ishan Durugkar, Josiah P. Hanna, and Peter Stone. Reducing Sampling Error in Batch Temporal Difference Learning. International Conference on Machine Learning (ICML), July 2020. Acceptance rate: 21.8%. [PDF].

Journal Articles

1. **Brahma S. Pavse***, Faraz Torabi*, Josiah P. Hanna, Garrett Warnell, Peter Stone. RIDM: Reinforced Inverse Dynamics Modeling for Learning From a Single Observed Demonstration. IEEE Robotics and Automation Letters, July 2020. **2nd place in the RoboCup 3D Sim Scientific Challenge 2019**.

Peer-reviewed Workshop Papers

3. Josiah P. Hanna, **Brahma S. Pavse**, and Abhinav Narayan Harish. Replacing Implicit Regression with Classification in Policy Gradient Reinforcement Learning. Workshop on Finding the Frame: An RLC Workshop for Examining Conceptual Frameworks, Reinforcement Learning Conference (RLC), August 2024.
2. **Brahma S. Pavse** and Josiah P. Hanna. Scaling Marginalized Importance Sampling to High-Dimensional State-Spaces via State Abstraction. Workshop on Offline Reinforcement Learning, Neural Information Processing Systems (NeurIPS), December 2022.
1. **Brahma S. Pavse**, Josiah P. Hanna, Ishan Durugkar, and Peter Stone. On Sampling Error in Batch Action-Value Prediction Algorithms. Workshop on Offline Reinforcement Learning, Neural Information Processing Systems (NeurIPS), December 2020.

Book Chapters

2. Patrick MacAlpine, Faraz Torabi, **Brahma Pavse**, and Peter Stone. UT Austin Villa: RoboCup 2019 3D Simulation League Competition and Technical Challenge Champions. In RoboCup 2019: Robot World Cup XXIII, Lecture Notes in Artificial Intelligence, Springer, 2019.

1. Patrick MacAlpine, Faraz Torabi, **Brahma Pavse**, John Sigmon, and Peter Stone. UT Austin Villa: RoboCup 2018 3D Simulation League Champions. In RoboCup 2018: Robot Soccer World Cup XXII, Lecture Notes in Artificial Intelligence, Springer, 2019.

Teaching Experience

University of Wisconsin – Madison, Madison, WI, USA

Teaching Assistant — Introduction to Artificial Intelligence **Fall 2024**

University of Texas at Austin, Austin, TX, USA

Teaching Assistant — Data Structures — Rating: 4.5/5.0 **Fall 2016**

Service

- Officer, Research to Impact at UW-Madison (2025).
- Coordinator, UW-Madison Reinforcement Learning Reading Group (2022-2025).
- Graduate Student Mentor, Wisconsin Science and Computing Emerging Research Stars [WISCERS] (2025, 2022).
- Reviewer, UT Austin Computer Science Dept. MS Admissions Committee (2020).

Reviewing

- Reinforcement Learning Conference (RLC) 2025, 2024.
- International Conference on Machine Learning (ICML) 2023-2025.
- International Conference on Learning Representations (ICLR) 2025, 2023, 2022.
- Neural Information Processing Systems (NeurIPS) 2022-2025.

- RLC Finding the Frame: An RLC Workshop for Examining Conceptual Frameworks 2024.
- NeurIPS Goal-Conditioned Reinforcement Learning Workshop 2023.
- Association for the Advancement of Artificial Intelligence (AAAI) 2023.
- International Conference on Robotics and Automation (ICRA) 2021.

Mentoring

UW Madison Undergraduates

- Stuti Pandey (2024-)
- Lucas Poon (2024). Next: CS PhD student at Oregon State University.
- Adhit Sankaran (2022 - 2023). Next: MS in CS at Cornell University.

Invited Talks

- Netflix Machine Learning and Inference seminar. July 2025.
- UW-Madison Systems, Information, Learning, Optimization (SILO) seminar. November 2024.
- UT Austin Reinforcement Learning Reading Group. April 2024.
- EdIntelligence at The University of Edinburgh. July 2020.

Technical Skills

- Languages: Python, Java, C++, Matlab
- Robotics simulators: MuJoCo
- Frameworks/Libraries/Tools: HTC Condor, PyTorch, OpenAI Gym, Pandas

Personal Details

- Citizenship: USA