Saaduddin Mahmud

+1 347-948-0507 | smahmud@umass.edu | saadmahmud.com

Education

University of Massachusetts Amherst — Massachusetts, USA
Ph.D. in Computer Science (4th Year) — *September 2021 – Present*

University of Massachusetts Amherst — Massachusetts, USA
M.Sc. in Computer Science (CGPA: 4.00/4.00) — *September 2021 – May 2024*

University of Dhaka — Dhaka, Bangladesh
B.Sc. in Computer Science and Engineering — *January 2016 – December 2019*

Experience

AI Research and Applications Researcher Intern — June 2024 – August 2024
Nissan Group North America, Inc — *Nissan Advanced Technology Center, Silicon Valley, USA*

- Effective methods to facilitate natural language communication between autonomous vehicles (AVs) and humans.

AI Research and Applications Researcher Intern — June 2023 – September 2023
Nissan Group North America, Inc — *Nissan Advanced Technology Center, Silicon Valley, USA*

- Explanation generation for autonomous vehicles (AVs).

Research Assistant — September 2021 – Present
Resource-Bounded Reasoning Lab, MCICS, UMass Amherst. — *Massachusetts, USA*

- Advisor: Professor Shlomo Zilberstein
- Reinforcement Learning and NLP

Patent

1. Vehicle Decision Making Using Sequential Information Probing.
   Under Review; Joint Inventor.

2. Dynamic Refinement of Custom Classes Using Zero-Shot Image
   Under Review; Joint Inventor.

Research Papers

1. Distributed Multi-Agent Coordination Using Multi-Modal Foundation Models
   Saaduddin Mahmud, Dorian Benhamou Goldfajn, and Shlomo Zilberstein. (Under Review).

2. MAPLE: A Framework for Active Preference Learning Guided by Large Language Models
   Saaduddin Mahmud, Meson Nakamura, and Shlomo Zilberstein. (**AAAI 2025**).

3. Verification and Validation of AI Systems Using Explanations
   Saaduddin Mahmud, Sandhya Saisubramanian, and Shlomo Zilberstein. (**AAAI Fall Symposium 2024**).

4. Explaining the Behavior of POMDP-based Agents Through the Impact of Counterfactual Information.
   Saaduddin Mahmud, Marcell VazquezChanlatte, Stefan Witwicki and Shlomo Zilberstein. (**AAMAS 2024**).

5. Learning Constraints on Autonomous Behavior from Proactive Feedback.
   Connor Basich*, **Saaduddin Mahmud***, and Shlomo Zilberstein. (**IROS 2023**).

6. Explanation-Guided Reward Alignment.
   Saaduddin Mahmud, Sandhya Saisubramanian, and Shlomo Zilberstein. (**IJCAI 2023**).

7. Semi-Autonomous Systems with Contextual Competence Awareness
   Saaduddin Mahmud, Connor Basich, and Shlomo Zilberstein. (**AAMAS 2023**).

8. Causal Explanations for Sequential Decision Making Under Uncertainty: Foundations and Analysis. Samer B. Nashed, **Saaduddin Mahmud**, Claudia V. Goldman, and Shlomo Zilberstein.(**Under Review JAIR, AAMAS 2023, Ext. Abs.**).

9. Estimating Causal Responsibility for Explaining Autonomous Behavior
Saaduddin Mahmud*, Samer B. Nashed*, Claudia V. Goldman, and Shlomo Zilberstein.
(**EXTRAAMAS workshop at AAMAS 2023**).

10. REVEALE: A Framework for Reward Verification and Learning.
Saaduddin Mahmud, Sandhya Saisubramanian, and Shlomo Zilberstein. (**SafeAI sy AAAI 2023, Best Paper Award Nomination**).

Honors & Awards

Distinguished PhD Candidate 2024
Recognition for outstanding performance in Ph.D. candidacy qualification.

B.Sc. Scholarship by the University Grants Commission 2021
Awarded for outstanding performance in B.Sc.

Programming Skills

Languages: Python, Julia, C, C++.
Libraries JAX, PyTorch.

Reference

Professor Shlomo Zilberstein | *Professor, CICS, University of Massachusetts Amherst, MA, USA*

- Email: shlomo@cs.umass.edu
- Personal Website: https://groups.cs.umass.edu/shlomo/