Sreejith Balakrishnan, Ph.D. Candidate, National University of Singapore

Singapore, +6597262108, sreejithb@gmail.com

LINKS

Google Scholar, LinkedIn

PROFILE

I am a Ph.D. candidate interested in developing Artificial Intelligence and Machine Learning algorithms with a focus on value alignment, fairness, ethics, and safety. My current research topics include value alignment through Inverse Reinforcement Learning, and fairness in algorithmic decision making. In addition to my experience as a researcher, I also have an industry career of 8 years. I have held many positions during my career including Research Engineer at the Singapore Management University and Section Manager at ABB Private Limited.

RESEARCH INTERESTS

Value Alignment

Algorithmic Fairness

Inverse Reinforcement Learning

Learning from Demonstrations

Causal inference

PUBLICATIONS

Dec 2020

Sreejith Balakrishnan, Quoc Phong Nguyen, Bryan Kian Hsiang Low, and Harold Soh. Efficient Exploration of Reward Functions in Inverse Reinforcement Learning via Bayesian Optimization. Neural Information Processing Systems (NeurIPS) 2020.

Aug 2022

Sreejith Balakrishnan, Jianxin Bi, and Harold Soh. SCALES: From Fairness Principles to Constrained Decision-Making. Artificial Intelligence, Ethics, and Society (AIES) 2022.

AWARDS

Jan 2021

Research Achievement Award, National University of Singapore

- Awarded for the NeurIPS 2020 accepted paper.

Jan 2019

Honor List of Student Tutors, National University of Singapore

- Awarded for outstanding tutoring for the module "AI Decision Making and Planning"

Aug 2016 — Jul 2020

NUS Research Scholarship, National University of Singapore

- Awarded to pursue Ph.D. program in School of Computing at National University of singapore.

EDUCATION

Aug 2016 — Present

Ph.D., National University of Singapore — Singapore

- Working on Inverse Reinforcement Learning and Value Alignment for robotics as well as Fairness in Decision Making.
- Relevant coursework: Uncertainty Modelling in AI, Advanced Topics in Artificial Intelligence

Aug 2010 — Jul 2013

Master of Science, National University of Singapore — Singapore

- M.Sc in Electrical Engineering with a specialization in Computer Engineering
- 4.6 GPA
- Relevant coursework: Real-time Systems, Linear Systems, Pattern Recognition, Evolutionary Computing

Aug 2004 — Jul 2008

Bachelor of Engineering, Nanyang Technological University — Singapore

- B.Eng in Electrical and Electronics Engineering with a Minor in Computing
- First Class Honors
- Relevant coursework: Robotics & Automation, Control Engineering Design, Modelling & control, Computer Vision

MACHINE LEARNING SKILLS

Inverse Reinforcement Learning	Expert	Deep Learning	Expert
Algorithmic Decision-making	Expert	Reinforcement Learning	Expert
Gaussian Process Regression	Expert	Active Learning	Experienced
Bayesian Optimization	Expert	Evolutionary Algorithms	Skillful

EMPLOYMENT HISTORY

Aug 2017 — Dec 2017

Teaching Assistant, AI Planning and Decision Making, School of Computing, NUS, Singapore

Singapore

- Conducted tutorials for the undergraduate (4th year) module
- Awarded Honor List of Student Tutors
- Received a high rating of 4.4 out of 5 based on student feedback.

Mar 2015 — Jul 2016

Research Engineer and Project Lead, School of Information Systems, SMU, Singapore

Singapore

- Technical lead for a Singapore urban planning project undertaken by Fujitsu-SMU Urban Computing & Engineering Corporate Lab (UNiCEN).
- Modeled the passenger flow at various taxi stands in Singapore to address the mismatch in taxi supply and demand.

Jun 2008 — Feb 2015

Section Manager, Regional After Sales Support, ABB Private Limited, Singapore

Singapore

- Spearheaded a team of 5 engineers to provide engineering solutions to customers in Asia, Australia, and New Zealand.
- Invited to conduct advanced-level product training for engineers and customers in various countries around the globe.

PROJECTS

Jun 2019 — Dec 2019

Decentralized Task Allocation for Multi-robot Systems, National University of Singapore

- Designed a task allocation algorithm for a decentralized multi-robot environments with limited communication.

Aug 2007 — Jul 2008

Control System for Humanoid Robot, Final Year Project, Nanyang Technological University, Singapore

- Interfaced a humanoid robot with sensors using various communication protocols like Bluetooth and I2C.

Jul 2007 — Dec 2007

Humanoid Robot Exhibition, Discovery Center, Singapore

- Led a team of 12 students to create a a public exhibition where the movements of 4 humanoid robots were synchronized against background music.

PROGRAMMING LANGUAGES

Python for machine learning	Highly proficient	Javascript for web development	Working knowledge
R for data analysis and visualization	Working knowledge	C/C#	Working knowledge
Java for desktop development	Working knowledge		