Amitayush Thakur

amitayusht@gmail.com | in/amitayush-thakur | [amit9oct](https://github.com/amit9oct) | amit9oct.github.io

Research Interests

Primary Interests: Interested in automated mathematical reasoning through LLMs and its implication in fully verified program synthesis essential for generating industry-grade code from AI.

Other Interests: Program Synthesis from natural language, Neuro-symbolic AI, Explainability of Neural Networks

Education

University of Texas at Austin — Austin, Texas

Ph.D. Computer Science — *Aug. 2022 – Present*

Advisor: Prof. Swarat Chaudhuri

Cumm. GPA - 4.00/4.00

Birla Institute Of Technology and Science, Pilani — Pilani, India

Master of Science (Hons.) Mathematics — *Aug. 2012 – July 2017*

Bachelor of Engineering (Hons.) Computer Science — *Aug. 2012 – July 2017*

Cumm. GPA - 9.02/10.00

Research Experience

Research Intern — Jan 2017 - June 2017

Microsoft Research — *Bangalore, India*

- Worked on Program Synthesis with Microsoft's *PROSE (Program Synthesis using Examples) SDK* team. Work was related to use of ML in PL
- It involved synthesizing programs which were generated on a restricted grammar using input-output examples, and then using ML to rank generated programs to get the most useful program.

Summer Research Fellow — May 2015 - July 2015

National Institute of Science Education & Research — *Bhubaneswar, India*

- Worked under Professor Brundaban Sahu in Computational Number Theory
- The main focus was studying the time complexity of Lenstra's Elliptic Curve Factorization Algorithm, and its relation to y-smooth numbers

Project Trainee — May 2014 - July 2014

Bhabha Atomic Research Centre (BARC) — *Mumbai, India*

- Worked in development of Network Management System and automation modules.

Software Engineering Experience

Software Engineer II — September 2019 – July 2022

Microsoft India Development Center — *Hyderabad, India*

Microsoft Azure Backup (MAB) team

- Developed technique for unsupervised pattern mining of service trace logs to predict root cause for the customer issues. The method was based on clustering similar logs and then training a decision tree on top of it to infer the reasoning behind the clustering. This inference was then used by service engineers to figure out the root cause for the customer issues. Mentored interns to implement the same.
- Developed & Designed a standalone micro-service to efficiently back up hierarchies of objects, such as File Shares, on Cloud. The micro-service split the entire backup into multiple smaller disjoint sub-tasks and then efficiently solved these sub-tasks in parallel—the idea was inspired from map reduce. Designed algorithms, along with proof of correctness, to figure out differences between two hierarchies of objects facilitating incremental backups i.e. only backing up the changed objects

Software Engineer — July 2017 - August 2019

Microsoft India Development Center — *Hyderabad, India*

Microsoft Azure Backup (MAB) team

- Focused on creating micro-services, distributed systems, using formal methods (TLA+), unit testing, scaling, Azure Cloud Storage, File Systems
- Developed and Designed Scalable & Distributed Billing orchestrator for Azure Backup. The overall time to compute billing related information improved **6x** because of the distributed design.

Software Engineering Intern May 2016 - July 2016

Microsoft India Development Center *Hyderabad, India*

Microsoft Azure Backup (MAB) team

- Improved performance for recovering parts of data from large (≥ 1 TB) backed-up disks, the performance went up from **10 hrs to 3 mins** using iSCSI and mocking certain layers of NTFS file system.

PUBLICATIONS

A. Gautam, A. Thakur, G. Dhanania and S. Mohan, ***"A distributed algorithm for balanced multi-robot task allocation,"*** 2016 11th International Conference on Industrial and Information Systems (ICIIS), Roorkee, 2016, pp. 622-627, doi: 10.1109/ICIINFS.2016.8263014. [IEEE Link]

OTHER SKILLS

Technologies: Git, Docker, Microsoft Azure Cloud, Distributed Systems, Map Reduce, Micro-Service Architecture, App Development, Web Development, Formal Methods, Deep Learning, RL
Languages: C#, Python, Java, C/C++, JavaScript
Libraries: PyTorch, Tensorflow, scikit-learn, pandas, NumPy, Matplotlib, spaCy