


AMITAYUSH THAKUR

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

RESEARCH INTERESTS

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

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

EDUCATION

University of Texas at Austin

Ph.D. Computer Science

Advisor: Prof. [Swarat Chaudhuri](#)

Austin, Texas

Aug. 2022 – Present

Birla Institute Of Technology and Science, Pilani

Master of Science (Hons.) Mathematics

Bachelor of Engineering (Hons.) Computer Science

Pilani, India

Aug. 2012 – July 2017

PUBLICATIONS & PREPRINTS

A. Thakur, J. Lee, G. Tsoukalas, M. Sistla, M. Zhao, S. Zetsche, G. Durrett, Y. Yue, and S. Chaudhuri, “**CLEVER: A Curated Benchmark for Formally Verified Code Generation**,” PrePrint, 2025. [[arXiv](#)][[Code](#)]

A. Thakur, G. Tsoukalas, G. Durrett, and S. Chaudhuri, “**PROOFWALA: Multilingual Proof Data Synthesis and Theorem-Proving**,” PrePrint, 2025. [[arXiv](#)][[Code](#)]

G. Tsoukalas, J. Lee, J. Jennings, J. Xin, M. Ding, M. Jennings, **A. Thakur**, and S. Chaudhuri. “**PutnamBench: Evaluating Neural Theorem-Provers on the Putnam Mathematical Competition**.” Neural Information Processing Systems (NeurIPS), 2024.(Best Paper Award at AI for Math Workshop ICML 2024) [[arXiv](#)][[Code](#)]

A. Thakur, G. Tsoukalas, Y. Wen, J. Xin, and S. Chaudhuri, “**An In-Context Learning Agent for Formal Theorem-Proving**,” First Conference on Language Modeling (COLM), 2024. [[arXiv](#)][[Code](#)]

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](#)]

INVITED TALKS

Special Seminar at Department of Mathematics at Rutgers, “**Advancing Theorem Proving with AI: The Power of Large Language Models**,” Feb 2024. [[Talk Details/Abstract](#)]

AWARDS

2024: Amazon Automated Reasoning Fellowship [UT-Amazon Science Hub](#)

EXPERIENCE

Software Engineer II

Microsoft India Development Center

Microsoft Azure Backup (MAB) team

September 2019 – July 2022

Hyderabad, India

- 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.

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.

OTHER SKILLS

Technologies: Git, Docker, Microsoft Azure Cloud, Distributed Systems, Map Reduce, Micro-Service Architecture, App Development, Web Development, Formal Methods, Theorem Proving, Deep Learning, RL

Languages: Python, C#, Lean, Rust, C/C++, Java, JavaScript

Libraries: PyTorch, ray, Tensorflow, scikit-learn, pandas, NumPy, Matplotlib, spaCy

Course Work: Neural Networks, Natural Language Processing, RL, Programming Languages, Compilers, Advanced Algorithms, Data Structure and Algorithms, OS, Cryptography, Number Theory, Real Analysis, Numerical Analysis, Graphs & Networks, Operations Research