Naman Jain

Website: naman-ntc.github.io ◇ Email: naman_jain@berkeley.edu

EDUCATION

University of California, Berkeley *2022 - present*
Ph.D. in Computer Science

Indian Institute of Technology, Bombay *2016 - 2020*
B.Tech (Honors) in Computer Science and Engineering

WORK AND INTERNSHIPS

Meta FAIR (Codegen), Seattle *May '24 - Dec '24*
Research Scientist Intern

AWS CodeWhisperer, Sunnyvale *June '23 - Aug '23*
Applied Research Intern

Microsoft Research (MSR), India *Aug '20 - Aug '22*
Research Fellow

PUBLICATIONS

19. GSO: SWE Agents Struggle at Reasoning and Engineering for Software Optimization
Manish Shetty, **Naman Jain**, Jinjian Liu, Vijay Kethanaboyina, Koushik Sen, Ion Stoica
Submitted to Neurips 2025, San Diego [preprint] [web]
18. R2E-Gym Procedural Environment Generation and Hybrid Verifiers for Scaling Open-Weights SWE Agents
Naman Jain*, Jaskirat Singh*, Manish Shetty, Liang Zheng, Koushik Sen, Ion Stoica
Submitted to COLM 2025, Canada [preprint] [web]
17. Challenges and Paths Towards AI for Software Engineering
Alex Gu, **Naman Jain***, Wen-Ding Li*, Manish Shetty*, Yijia Shao, Ziyang Li, Diyi Yang, Kevin Ellis, Koushik Sen, Armando Solar-Lezama
Proceedings of ICML 2025, Canada [preprint]
16. CopilotArena: A Platform for Code LLM Evaluation in the Wild
Wayne Chi, Valerie Chen, Anastasios Nikolas Angelopoulos, Wei-Lin Chiang, Aditya Mittal, **Naman Jain**, Tianjun Zhang, Ion Stoica, Chris Donahue, Ameet Talwalkar
Proceedings of ICML 2025, Canada [preprint]
15. Syzygy: Dual Code-Test C to Rust Translation using LLMs and Dynamic Analysis
Manish Shetty*, **Naman Jain***, Adwait Godbole*, Sanjit Seshia, and Koushik Sen
Proceedings of LLM4Code@ICSE 2025, Ottawa, Canada [preprint]
14. LiveCodeBench: Holistic Contamination Free Evaluation of Code LLMs
Naman Jain, King Han, Alex Gu*, Wen-Ding Li*, Fanjia Yan*, Tianjun Zhang*, Sida Wang, Armando Solar Lezamma, Koushik Sen and Ion Stoica
Proceedings of ICLR 2025, Singapore [preprint]
13. BigCodeBench: Benchmarking Code Generation with Diverse Function Calls and Complex Instructions
Zhou et. al. *Proceedings of ICLR 2025, Singapore* [preprint]
12. R2E: Turning Any Github Repository to Programming Agent Test Environment
Naman Jain*, Manish Shetty*, Tianjun Zhang, King Han, Koushik Sen and Ion Stoica
Proceedings of ICML 2024, Vienna, Austria [paper]

11. RAFT: Adapting Language Model to Domain Specific RAG
Tianjun Zhang, Shishir G. Patil, **Naman Jain**, Sheng Shen, Matei Zaharia, Ion Stoica and Joseph E. Gonzalez
Accepted at COLM 2024 [paper]
10. Selfcodealign: Self-alignment for code generation
Yuxiang Wei, Federico Cassano, Jiawei Liu, Yifeng Ding, **Naman Jain**, Zachary Mueller, Harm de Vries, Leandro Von Werra, Arjun Guha, Lingming Zhang
Proceedings of Neurips 2024, Vancouver, Canada [paper]
9. The Counterfeit Conundrum: Can Code Models Grasp the Nuances of Their Incorrect Generations?
Alex Gu, Wen-Ding Li*, **Naman Jain***, Theo X. Olausson*, Celine Lee*, Koushik Sen, and Armando Solar-Lezama
Submitted to ACL 2024, Bangkok, Thailand [preprint]
8. StarCoder 2 and The Stack v2: The Next Generation
Lozhkov et. al.
To be submitted to TMLR 2024 [preprint]
7. LLM-Assisted Code Cleaning For Training Accurate Code Generators
Naman Jain, Tianjun Zhang, Wei-Lin Chiang, Joseph E. Gonzalez, Koushik Sen and Ion Stoica
Proceedings of ICLR 2024, Vienna, Austria [paper]
6. Revisiting Prompt Engineering via Declarative Crowdsourcing
Aditya G Parameswaran, Shreya Shankar, Parth Asawa, **Naman Jain** and Yujie Wang
Proceedings of ICDE 2023, Anaheim, California [preprint]
5. StaticFixer: From Static Analysis to Static Repair
Naman Jain, Shubham Gandhi, Atharv Sonwane, Aditya Kanade, Nagarajan Natarajan, Suresh Parthasarathy, Sriram Rajamani and Rahul Sharma
Preprint Available [preprint]
4. Jigsaw: Large Language Models meet Program Synthesis
Naman Jain, Skanda Vaidyanath, Arun Iyer, Nagarajan Natarajan, Suresh Parthasarathy, Sriram Rajamani and Rahul Sharma
Proceedings of ICSE 2022, Pittsburgh, Pennsylvania [paper]
3. Learning Accurate Decision Trees with Bandit Feedback via Quantized Gradient Descent
Ajaykrishna Karthikeyan*, **Naman Jain***, Nagarajan Natarajan, and Prateek Jain
Proceedings of TMLR 2022 [paper]
2. What's in a Name? Are BERT Named Entity Representations just as Good for any other Name?
Sriram Balasubramanian*, **Naman Jain***, Gaurav Jindal, Abhijeet Awasthi and Sunita Sarawagi
Workshop Proceedings of ACL 2020, Virtual Conference [paper] [supplement]
1. On the Robustness of Human Pose Estimation
Naman Jain*, Sahil Shah*, Abhishek Sharma and Arjun Jain
Workshop Proceedings of CVPR 2019, Long Beach, California [paper] [supplement]

* equal contribution

RESEARCH PROJECTS

Long-Horizon Code Evaluation Aug'24 - Present

· Semi-autonomous construction of benchmark for complex long-horizon coding problems
· Curating like optimization, translation, refactoring, bug-finding, and repair

LiveCodeBench - Contamination Free and Holistic Code Evaluation Oct'23 - Dec'23

· Performing periodically updated "live" benchmarking for code LLMs to evade data contamination
· Holistically evaluating LLMs for code generation, repair, execution, optimization, and test generation

CodeRefuse - Model Refusal for Coding Assistants Jun'23 - Aug'23

- Constructed a high-quality model refusal dataset for user-queries using retrieval and in-context learning
- Transformed CWE docs from OWASP and MITRE into unsafe requests for data annotation pipeline

Parametric Template Distillation for API Summarization Jan'22 - Apr'22

- Learning Parametric Templates (natural language with programmatic holes) for API summarization

Static Repair of Information Flow Vulnerabilities Sep'21 - Nov'22

- Proposed a novel static-analysis-based approach for mining high-quality program-repair edits
- Repaired cross-site-scripting and unvalidated-call vulnerabilities in over 1000 repositories (90% recall)

Jigsaw – Combining Language Models with Program Synthsis Dec'20 - Sep'21

- Proposed architecture for augmenting black-box models with program analysis and synthesis-based post-processing block that provides correctness guarantees and also allows learning from user feedback

SELECTED SOFTWARE AND OPEN SOURCE

- Human Pose Estimation - PyTorch (stats - ★>450, ᛘ>70)
- Implemented `torch.randint` in PyTorch! (open source contribution)

SCHOLASTIC ACHIEVEMENTS

- Awarded Undergraduate Research Award (**URA**) for Autumn 2018 (2 out of 121 students)
- Secured **All India Rank 36** in JEE Advanced 2016 among **0.15 million** candidates
- Awarded Gold Medal for being among top **35** students in Indian National Physics Olympiad (**INPhO**)
- Among the **top 300** in Indian National Chemistry and Astronomy Olympiads (**INChO** & **INAO**)
- Awarded KVPY Fellowship and NTSE Scholarship by Govt. of India

RELEVANT COURSES

- **Computer Science -** Automata Theory, Compilers, Parallelizing Compilers, Advanced Machine Learning, Learning Agents (RL), Computer Vision, Data Structures, Algorithms, Operating System
- **Mathematics -** Calculus, Linear Algebra, Differential Equations, Numerical Analysis, Discrete Maths

TECHNICAL STRENGTHS

Strong	Python, C, C++, C#, Racket, Prolog, Bash
Web	Flask, ASP.NET, Guincorn, Nginx, IIS, WebSockets, React, JSP
Tools	Git, LaTeX, OpenGL, OpenMP, Gnuplot, Doxygen

TEACHING & MENTORSHIP

- Teaching Assistant for the course Automatic Speech Recognition under Prof. Preethi Jyothi
- Teaching Assistant for the course Quantum Mechanics under Prof. Aftab Alam

REFERENCES

Prof. Koushik Sen
University of California, Berkeley
webpage ◇ email

Prof. Ion Stoica
University of California, Berkeley
webpage ◇ email

Dr. Sriram Rajamani
Microsoft Research, India
webpage ◇ email

Dr. Prateek Jain
Google Research, India
webpage ◇ email

Prof. Sunita Sarawagi
Indian Institute of Technology, Bombay
webpage ◇ email

Dr. Nagarajan Natarajan
Microsoft Research, India
webpage ◇ email