Aman Goel

amangoel.ai | aman@amangoel.ai | Scholar | GitHub | LinkedIn

Research Focus

I spent a decade building automated provers — a competition-winning model checker (**1st, HWMCC 2020**) and the **first automatic proof of Lamport's Paxos** — and now apply that rigor to LLM safety. At AWS, I shipped the neurosymbolic guardrails behind Amazon Bedrock's Automated Reasoning checks (**CAV 2026**, **99.2% soundness**), and built zero-knowledge hallucination detection (**EMNLP 2025**: **39% F1 gain**, **+9pp** on GPQA-diamond via cross-model test-time compute) and RL-guided red teaming (**NAACL 2025 oral**: **≥95% ASR** on GPT-4o/GPT-4 Turbo, 3× fewer queries) deployed inside AWS. My thesis: **LLMs can't be trusted without proofs** — my work maps what formal verification can and cannot do about that. **17+ papers** (CAV, EMNLP, NAACL, SOSP, ATC, FMCAD); prototype to production in months.

Experience

Amazon Web Services – Automated Reasoning Group — Seattle, WA

Senior Applied Scientist | Neurosymbolic AI — *Nov 2021 – Present*

- **Built Neurosymbolic Guardrails** | *Shipped as Automated Reasoning checks in Amazon Bedrock Guardrails · **CAV 2026***
- Co-built the pipeline: LLM formalizes policies → SMT solver verifies claims → **99.2% soundness**; catches hallucinations pure-LLM approaches miss; generally available since Aug 2025, serving production traffic
- **Invented Zero-Knowledge Hallucination Detection** | *EMNLP 2025*
- Designed cross-model consistency framework requiring no external knowledge; **39% F1 improvement** over baselines; **9pp accuracy gain** on GPQA-diamond; rigorous evaluation on 5 benchmarks, 8 models
- **Built Automated Red Teaming System** | *NAACL 2025 Oral · Open-sourced*
- Engineered mutation-based fuzzer with RL-guided selection; **≥95% ASR** on GPT-4o/GPT-4 Turbo in **3× fewer queries**; deployed for internal red teaming
- **Owned Distributed Systems Verification** | *Aurora, DynamoDB, MemoryDB*
- Built systematic fuzzing + model checking tools; found critical design bugs missed by testing; drove adoption across 10+ AWS teams; reduced verification time from weeks to hours
- **Research leadership**: mentored 5+ scientists & engineers, supervised 4+ intern projects; PC member, **CAV 2025**; sustained independent formal-methods agenda alongside product work (CAV'25, ISoLA'24, FORTE'23, FMCAD'23)

University of Michigan — Ann Arbor, MI

PhD Researcher | Advisor: Karem Sakallah — *Aug 2016 – Oct 2021*

- **IC3PO: First Automatic Proof of Lamport's Paxos** | *FMCAD 2021*
- Invented symmetry-aware IC3 algorithm for infinite-state verification; enabled Sift (ATC'22), Bakery proof (FORTE'23)
- **AVR: Competition-Winning Model Checker** | *1st HWMCC 2020 · medals at HWMCC 2024 & 2025*
- Engineered word-level abstraction-refinement; **7 gold medals**; benchmarked on 1000+ industrial designs

SRI International (intern) — **RL-guided quantifier instantiation** in the Yices 2 SMT solver — Summer 2020

Cadence (research intern) — word-level verification engines for JasperGold; shipped to production — Summer 2019

Education

University of Michigan — Ph.D., Computer Science · advisor: Karem Sakallah — Aug 2016 – Oct 2021

Thesis: *From Finite to Infinite: Scalable Automatic Verification of Hardware Designs and Distributed Protocols*

IIT Madras — B.Tech EE + M.Tech Microelectronics · GPA 9.23/10, **Silver Medalist** — Jul 2011 – May 2016

Selected Publications

A Neurosymbolic Approach to Natural Language Formalization and Verification – **CAV 2026**
C. An, S. Bayless, et al. (30 authors, alphabetical, incl. ***A. Goel****) · to appear · deployed in Amazon Bedrock Guardrails*

Zero-Knowledge LLM Hallucination Detection via Cross-Model Consistency – **EMNLP 2025**
A. Goel*, D. Schwartz, Y. Qi*

TurboFuzzLLM: Mutation-based Fuzzing for Jailbreaking LLMs – **NAACL 2025** *Oral*
A. Goel*, X. Wu, Z. Wang, D. Bespalov, Y. Qi*

Sift: Refinement-guided Automation for Distributed Systems – **USENIX ATC 2022**
H. Ma, H. Ahmad, ***A. Goel****, et al.*

Towards an Automatic Proof of Lamport's Paxos – **FMCAD 2021**
A. Goel*, K. Sakallah*

I4: Incremental Inference of Inductive Invariants – **SOSP 2019**
H. Ma, ***A. Goel****, et al.*

17+ publications at ***CAV, EMNLP, NAACL, SOSP****, USENIX ATC, FMCAD, TACAS, NFM, DATE*

Awards & Recognition

1st Place, Hardware Model Checking Competition 2020 – *7 Gold, 1 Silver, 1 Bronze* (vs. industry tools)
Medals, HWMCC 2024 & 2025 – *gold + silver (2024) · silver, word-level arrays (2025)*
Rackham Predoctoral Fellowship 2020 – *Top PhD research at Michigan* · Simons Institute visitor (2021)
Best Student Research Award – *CSE Graduate Honors Competition 2019*
Silver Medalist – *IIT Madras, Branch Rank 2/120 in Electrical Engineering*
National Award for Empowerment of Persons with Disabilities – *Govt. of India, 2013*

Technical Skills

Neurosymbolic & Formal: SMT solvers (Z3, Yices), Lean, IC3/model checking, NL-to-formal-spec translation, proof automation

LLM Safety: hallucination detection, jailbreak fuzzing & red teaming, guardrails, test-time compute (MCTS, cross-model consistency), LLM-as-judge evaluation

Engineering: Python, C++, C, Java | PyTorch | AWS (Bedrock, SageMaker) | large-scale evaluation pipelines

Open Source & Research Tools

TurboFuzzLLM – LLM red-teaming fuzzer (NAACL 2025 oral); RL-guided mutation fuzzing	Python
AVR – Hardware model checker; HWMCC 2020 winner; used by Cadence, academia	C++, Python
IC3PO – Protocol verifier; first automatic proof of Lamport's Paxos (to our knowledge)	Python
Contributor: **Yices 2** (SMT solver), **Yosys** (synthesis)	C, C++

Service & Leadership

Program Committee: CAV 2025 (top formal methods venue), FMCAD 2022

Artifact Evaluation: SOSP 2023/2021, CAV 2023/2020, OSDI 2021, MICRO 2021

Mentorship: Mentored 5+ scientists & engineers; supervised 4+ intern projects on formal verification and AI safety

Press: Hacker News front page (159 points, 68 comments) · Michigan Engineering features · AWS News Blog, Automated Reasoning checks GA

Talks: FMCAD 2021 Paxos talk (recorded) · Panelist, ACL 2026 mentorship session on AI tools & career growth