Autoformalization and the New Era of Mathematical Discovery: Insights from Neel Somani

Autoformalization and the New Era of Mathematical Discovery: Insights from Neel Somani

Berkeley, California Feb 11, 2026 (Issuewire.com**)** - The intersection of artificial intelligence and pure mathematics is rapidly evolving, moving beyond simple computation into the realm of creative discovery. Recently, Neel Somani, a Berkeley-educated computer scientist and the founder of the blockchain platform Eclipse, spearheaded an experiment that offers a glimpse into this future. By organizing a group of undergraduate students to deploy advanced AI models against open Erdős problems, Somani did not just seek to solve math problems; he sought to understand the architecture of discovery itself.

The experiment, dubbed "GPT-Erdos," utilized tools like GPT-5.2 Pro and Deep Research to attack unsolved mathematical conjectures. The quantitative results were impressive—yielding accepted solutions, partial results, and undocumented rediscoveries. However, according to Somani, the true value of this exercise was not the raw output, but what the process revealed about the hidden, informal

concepts that guide human research. As AI begins to engage in "autoformalization"—converting human-readable proofs into machine-checkable formats—it is forcing the scientific community to confront the nuances of novelty, progress, and correctness.

The Ambiguity of Underspecification

One of the most profound findings from Somani's research involves the concept of underspecification. When an AI generates a solution, it often exposes the lack of precision in how humans define success. During the GPT-Erdos experiment, Somani observed instances where the AI produced a solution that solved the problem as stated, but it heavily repurposed existing results.

This raises a critical question regarding categorization: Is such a result a novel discovery, a mere rediscovery, or an extension of previous work? Neel Somani notes that disagreements over novelty are not purely epistemic; they serve as a proxy for intellectual contribution. When an AI produces a solution without a clear historical lineage, the human desire for a clean delineation of "novelty" clashes with the messy reality of mathematical derivation. The experiment highlighted that failure modes in AI research are often not result-based errors, but rather failures of specification—instances where the AI succeeds technically, but fails to satisfy the ill-defined human criteria of what constitutes a "new" contribution.

Defining Novelty in an Automated Age

The challenge of defining novelty is not limited to machines; it divides top mathematicians as well. Neel Somani points to instances where leading figures in the field, such as Terence Tao, might classify an AI-generated result as novel, while others might view it as derivative of existing theorems. This discord suggests that the mathematical community relies on intuition rather than formal logic to determine the value of a proof.

Neel Somani proposes that the industry may need to move toward a formal definition of novelty. This could potentially be framed around the minimum complexity required to express a proof. If a proof is simply an existing theorem with new parameters, it lacks novelty. However, if a proof necessitates the construction of several new non-trivial theorems that cannot be bypassed, it likely represents a genuine advancement. Drawing on his background in quantitative research and cryptography, Somani suggests drawing inspiration from zero-knowledge proofs, defining mathematical "knowledge" as the ability to reconstruct a proof using existing results in polynomial time.

The Problem of "Interestingness"

Beyond the mechanics of proving theorems, there lies a deeper, more abstract challenge: identifying which problems are worth solving. Human mathematicians possess an innate sense of "interestingness"—a heuristic that balances difficulty with utility. Large Language Models (LLMs), however, lack this intuition. They do not inherently understand which mathematical inquiries might unlock solutions in physics or engineering, nor do they grasp the cultural or aesthetic weight of a problem.

Neel Somani argues that this limitation extends beyond mathematics into business and art. Just as an AI struggles to discern which business ideas are genuinely novel, it struggles to identify meaningful mathematical pursuits. These values are not part of the training data. Consequently, the rise of autoformalization acts as a mirror, revealing that the "soft" concepts humans take for granted are actually the invisible guardrails of progress.

From Pure Math to Reliable Software

While the philosophical implications of autoformalization are vast, the practical applications are immediate, particularly in fields requiring absolute reliability. Somani, whose work with Eclipse focuses on decentralized technology, sees a direct line between formal math proofs and software security. In quantitative finance and blockchain development, the goal is often to create models and systems that are provably correct.

The sheer volume of code currently being generated by AI assistants poses a new risk profile. Somani refers to this as "slop code"—software produced so rapidly that human review becomes a bottleneck. Autoformalization offers a solution by applying formal methods at scale. Rather than relying on human oversight to catch memory safety violations in C++ kernels or exception handling errors, formalized AI systems could provide provable guarantees. This shift would make formal verification, once considered too cumbersome for general software development, a viable standard for critical infrastructure.

The Search for a "Closeness" Metric

Looking toward the future of this technology, Neel Somani identifies a significant gap in the current toolkit: the lack of a metric for "closeness" to completion. Current formal verification is binary; a proof either verifies, or it does not. However, the history of scientific discovery is rarely so black and white. Major breakthroughs, such as the Einstein field equations, were often discovered via heuristics and metaphors long before they were cleanly formalized.

Somani envisions a future where autoformalization includes a differentiable surrogate function that can measure how close a proof is to being correct. This would allow researchers to differentiate between a proof that is fatally flawed and one that is merely a few steps away from verification. Such a development would transform AI from a binary checker into a true collaborator, capable of navigating the heuristic, messy process of discovering truth.

Shaping the Future of Inquiry

The experiment led by Neel Somani suggests that even if AI progress were to stall today, the practice of mathematics has already been fundamentally altered. The ability to verify proofs via machine and quickly assimilate existing approaches allows researchers to bypass the rote memorization of literature and focus on high-level conceptualization.

As the founder of Eclipse and a mentor to the next generation of computer scientists, Somani continues to explore how these technologies can reshape decentralized systems and academic inquiry alike. The future of math is not just about machines solving problems; it is about machines helping humans understand the very nature of the problems they wish to solve.

To learn more visit: https://www.neelsomaniblog.com/

Media Contact

businessnews@mail.com

*********@gmail.com

Source : Neel Somani

See on IssueWire