AI News Digest

11/09/2026 - 11/15/2026

WEEKLY DIGEST: DATA & MATHEMATICAL SCIENCES

AI Powers New Era of Formal Verification and Proof

Google DeepMind unveiled AlphaProof Code, a multi-LLM agentic system that can formally verify compiled software and mathematically prove critical security properties. In parallel, Microsoft Research announced the resolution of the 31-year-old KLS conjecture, highlighting AI-assisted techniques in deep combinatorial proof work. A practical tutorial demonstrated how to construct a three-state Markov chain in Python, showing how AI tools are lowering the barrier to stochastic modeling. Together, these stories illustrate a rapid convergence of AI and rigorous mathematics, reshaping both software safety and theoretical research.

KEY FINDINGS

✓ AlphaProof Code combines multiple LLM agents with human expertise to formally verify compiled software and prove key security properties.

✓ Researchers announced a proof that finally resolves the 31-year-old KLS conjecture, a landmark result in combinatorial mathematics.

✓ A step-by-step guide shows how to build a three-state Markov chain in Python using NumPy, exemplifying AI-driven educational content for stochastic modeling.

IMPLICATIONS

→ AI-driven verification frameworks could become standard for certifying safety-critical software across industries.

→ The success on the KLS conjecture suggests AI-augmented proof assistants may accelerate the resolution of other long-standing mathematical problems.

→ Accessible tutorials on probabilistic models will broaden the skill set of data scientists, fostering more widespread use of Markov processes in research and industry.

> Key Metrics: 31-year-old KLS conjecture finally proved; Three mutually exclusive states (drafting, revising, distracted) used in the Markov chain example; AlphaProof Code leverages a multi-LLM architecture (multiple language models collaborating) to achieve formal verification

1. Introducing AlphaProof Code: a breakthrough in verifying complex software at depth

Category: Research
Departments: Computer Science, Mathematical Sciences
Topics: Enterprise AI, Multi-LLM Agentic Framework, LLM-Driven Formal Verification, Lean Code Generation for Security Proofs

Impact Score: 27.0
Rubric Scores

Research Relevance: 5	Educational Value: 4	Innovation / Novelty: 5
Practical Impact: 5	Interdisciplinary Potential: 5	Ethical / Policy Implications: 3

Google DeepMind introduced AlphaProof Code, a multi-LLM agentic framework that combines human expertise with AI agents to perform formal verification of software. Using this system, the researchers mathematically proved key security properties of the compiled Signal cryptographic protocol, generating over six million lines of Lean code and requiring human review of only about one percent of the codebase. The verification, which covered the actual implementation rather than just the design, was completed in four weeks-a process that traditionally takes expert teams years. Validation tests showed AlphaProof Code correctly identified two previously fixed bugs as insecure, confirming its ability to distinguish correct from buggy code.

https://deepmind.google/blog/introducing-alphaproof-code/

2. Microsoft Research Focus - Edition #98

Category: Research
Departments: Computer Science, Mathematical Sciences
Topics: Natural Language Processing, Latent Core Tokenizer, Context Compaction Benchmark, APEC Summarization

Impact Score: 20.0
Rubric Scores

Research Relevance: 4	Educational Value: 3	Innovation / Novelty: 4
Practical Impact: 4	Interdisciplinary Potential: 4	Ethical / Policy Implications: 1

Microsoft Applied Science Manager Jinyu Li discussed recent AI developments that move universal translation closer to practical implementation in the "Our Digital Life" podcast. Researchers announced a proof resolving the 31-year-old KLS conjecture, establishing that flat cuts identify bottlenecks of log-concave distributions within a fixed factor in any dimension, with the proof independently validated and formally verified. A study of 13.5 million coding sessions led to COMPACTIONBENCH, a benchmark for context compaction, and the APEC method, which prepares summaries in the background, achieving up to 25.4 % cost reduction and 9.6 % faster execution for AI coding agents. Researchers introduced the Latent Core Tokenizer (LCT), which extracts reusable word components before building a shared multilingual vocabulary, yielding improved tokenization efficiency across 104 languages.

https://www.microsoft.com/en-us/research/podcast/microsoft-research-focus-edition-98/

3. How to Build a Markov Chain in Python With NumPy

Category: Education
Departments: Computer Science, Mathematical Sciences
Topics: Reinforcement Learning, Markov Chain Modeling, Stationary Distribution Estimation, Memoryless Assumption Limitations

Impact Score: 10.0
Rubric Scores

Research Relevance: 1	Educational Value: 3	Innovation / Novelty: 1
Practical Impact: 2	Interdisciplinary Potential: 3	Ethical / Policy Implications: 0

The article introduces Markov chains as a mathematical framework for modeling a writer's daily activity, defining three mutually exclusive states-drafting, revising, and distracted-and presenting a transition matrix with specific probabilities for moving between these states. It explains the memoryless Markov property, demonstrates how to compute multi-step probabilities and stationary distributions, and provides a Python implementation of the model. The piece also cautions that, despite the ability to represent a process as a Markov chain, the assumption that the current state fully captures relevant past information may not hold for real-world behaviors.

https://towardsdatascience.com/how-to-build-a-markov-chain-in-python-with-numpy/