THE POTENCY OF AI IN SOLVING COMPLEX MATHEMATICAL PROBLEMS: THE PROSPECT AND STRATEGIES FOR MATHEMATICS

By

Nathan, Ndifreke Adolf,
Department of Science Education
Akwa Ibom State University, Ikot Akpaden,
Mkpat Enin l. G. Area

And

Samuel Matthew Akpan, Ph.D
(Senior Lecturer, Measurement and Evaluation, Research and Statistics.)
Department of Educational Foundations,
University of Calabar,
Calabar, Nigeria

ABSTRACT

This paper explores the transformative role of Artificial Intelligence (AI) in addressing complex mathematical problems, highlighting its growing potency, strategic approaches, and long-term prospects. With the advancement of deep learning, symbolic reasoning, and reinforcement learning, AI systems have demonstrated capabilities in solving differential equations, performing symbolic integration, and generating formal proofs. The study discusses emerging strategies such as neuro-symbolic models and automated theorem proving, which combine human-like reasoning with computational power. Additionally, it reviews landmark achievements where AI has contributed to new mathematical insights through collaboration with experts. The prospects of AI in mathematics indicate enhanced efficiency, creativity, and accessibility in problem-solving. This work also stresses the importance of ethical deployment, transparency, and interdisciplinary collaboration. Strategic integration of AI in education and research will further unlock its potential. By augmenting human intuition and logical thinking, AI is set to redefine the landscape of mathematical innovation. The study concluded that As AI tools evolve, they promise to enhance creativity, not just computation. This synergy is reshaping mathematical research and problem-solving. The study also recommended that governments and academic bodies should fund research into hybrid AI models that combine symbolic logic with deep learning to advance AI's reasoning capabilities in mathematics.

KEYWORDS: Artificial Intelligence, Complex Mathematical Problems and Education.

INTRODUCTION

The convergence of AI with symbolic reasoning and neural networks has given rise to neuro-symbolic approaches—systems that combine statistical learning with logical reasoning. These hybrid models are particularly potent in mathematical tasks that demand both abstraction and structure. Choi et al. (2021) emphasized how such systems are transforming the reasoning landscape by enabling AI to perform deductive tasks, generate proofs, and manipulate complex expressions (Choi et al., 2021). These strategies open new avenues in both automated theorem proving and the generation of conjectures, allowing AI to transition from a computational assistant to an intelligent collaborator in mathematical discovery.

AI's promise in mathematical innovation is further evidenced by its success in discovering new patterns and supporting intuitive reasoning. In an exemplary collaboration, Davies et al. (2021) reported how Deep Mind's AI guided mathematicians in uncovering new structures in knot theory and representation theory, thus supporting the development of novel mathematical conjectures (Davies et al., 2021). This capability illustrates AI's potential not just in validating existing knowledge, but in guiding theoretical exploration—a profound shift in the methodology of mathematical research that signals a future where human intuition is augmented by computational creativity.

Strategically, the development of AI-driven problem-solving systems relies on several key methodologies, including reinforcement learning, curriculum learning, and graph-based neural networks. These strategies equip AI to navigate complex solution spaces and adaptively learn mathematical operations. The Holist system developed by Alemi et al. (2021) applied reinforcement learning to improve theorem proving in higher-order logic, showing substantial performance gains in formal proof environments (Alemi et al., 2021). Such advances suggest that with the right strategies, AI can continue to deepen its involvement in mathematical problem-solving, offering both immediate utility and long-term prospects for innovation.

CONCEPT OF ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) is the simulation of human intelligence processes by machines, particularly computer systems, enabling them to perform tasks that typically require human intelligence, such as learning, reasoning, and problem-solving. Craig, Laskowski, & Tucci, (2024) Artificial intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems. Examples of AI applications include expert systems, natural language processing (NLP), and speech recognition and machine vision. As the hype around AI has accelerated, vendors have scrambled to promote how their products and services incorporate it. Often, what they refer to as "AI" is a well-established technology such as machine learning

Artificial intelligence (AI) is the theory and development of computer systems capable of performing tasks that historically required human intelligence, such as recognizing speech, making decisions, and identifying patterns. AI is an umbrella term that encompasses a wide variety of technologies, including machine learning, deep learning, and natural language processing (NLP). According to Staff, (2024), Yet, despite the many philosophical disagreements over whether "true" intelligent machines actually exist, when most people use the term AI today, they're referring to a suite of machine learning-powered technologies, such as Chat GPT or computer vision, that enable machines to perform tasks that previously only humans can do like generating written content, steering a car, or analyzing data.

Although the term is commonly used to describe a range of different technologies in use today, many disagree on whether these actually constitute artificial intelligence. Instead, some argue that much of the technology used in the real world today actually constitutes highly advanced machine learning that is simply a first step towards true artificial intelligence, or "general artificial intelligence" (GAI). Stryker, (2024) Artificial intelligence (AI) is technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy.

CONCEPT OF COMPLEX MATHEMATICAL PROBLEMS

Complex mathematical problems involve multiple steps, concepts, and operations, often requiring a deeper understanding of mathematical principles and the ability to apply them strategically. They can encompass various mathematical domains and challenge individuals to think critically and creatively (Vorhölter & Leiss, 2019). Complex problems rarely have straightforward, one-step solutions. They often require a series of calculations and manipulations to arrive at the answer. Understanding complex math problems is crucial not just for students and teachers, but also for anyone applying math in daily life or at work. Breaking down these challenges into manageable steps makes them much less intimidating (Vye et al., 2024). Solving the challenge involves formulating sub problems, organizing these subproblems into solution plans, differentiating solution-relevant from solution-irrelevant data, coordinating relevant data with appropriate subproblems, executing computations, and deciding among alternative solutions.

According to Kaitera and Harmoinen, (2022), Mathematical problem-solving requires skills to apply variety of different solution strategies and models. It is not uncommon that while students may excel on routine exercises (those that they have already seen and practiced), they fail to solve problems that differ from those they have previously encountered. Many mathematical problems have been stated but not yet solved. These problems come from many areas of mathematics, such as theoretical physics, computer science, algebra, analysis, combinatorics, algebraic, differential, discrete and Euclidean geometries, graph theory, group theory, model theory, number theory, set theory, Ramsey theory, dynamical systems, and partial differential equations.

THE IMPACT OF AI IN SOLVING COMPLEX MATHEMATICAL PROBLEMS

The impact of Artificial Intelligence (AI) on solving complex mathematical problems has grown exponentially since 2020. With advances in deep learning, symbolic computation, and reinforcement learning, AI has evolved from being a computational assistant to a semi-autonomous problem solver in mathematics. One of the landmark achievements was demonstrated by Lample and Charton (2020), who used transformer models to tackle symbolic integration and differential equations with performance comparable to traditional systems like Mathematica and Maple. Their work, published in Nature, marked a significant turning point by proving that neural networks could handle symbolic reasoning without explicit programming rules (Lample & Charton, 2020).

AI has also begun to show potential in discovering new mathematical patterns and insights. In a collaborative effort between DeepMind and mathematicians, Davies et al. (2021) used AI to guide intuition in pure mathematical fields like representation theory and knot theory. By applying deep learning to large mathematical datasets, their AI system uncovered previously unnoticed patterns, assisting researchers in formulating new conjectures. This breakthrough, also published in Nature, demonstrates AI's potential not only in solving equations but in augmenting human creativity in abstract mathematical reasoning (Davies et al., 2021).

Equally transformative has been the application of AI in formal theorem proving. Researchers like Polu and Sutskever (2020) introduced GPT-like models trained on formal logical systems such as Metamath to autonomously generate and prove mathematical theorems. These models significantly accelerate the proof discovery process and help verify logical consistency at scale. More recent developments, such as Alemi et al.'s (2021) work using Graph Neural Networks (GNNs) for proof search, have further improved the accuracy and efficiency of theorem proving within large formal libraries like HOL Light (Polu & Sutskever, 2020; Alemi et al., 2021).

In addition to theorem proving, AI has impacted the structuring of learning processes in mathematics. Tian et al. (2022) proposed curriculum learning strategies that gradually expose AI systems to more complex formal mathematical statements, improving the models' ability to generalize across domains. This approach mimics how humans learn mathematics, starting from simple axioms and building up to advanced concepts. Zheng et al. (2021) also introduced the minif2f benchmark, a dataset derived from Lean that allows for more precise evaluation of formal reasoning capabilities in neural models (Tian et al., 2022; Zheng et al., 2021).

A crucial contribution of AI lies in its capacity to blend neural networks with symbolic reasoning, creating hybrid systems that leverage the strengths of both. Chaudhuri and Cheng (2020) reviewed these neural-symbolic systems, highlighting their ability to bridge the gap between statistical pattern recognition and logical inference. These systems are particularly suited for mathematical problem-solving, where both intuitive approximation and exact logic are necessary. Recent studies like those by Joshi et al. (2022) have explored architectures such as skip-tree networks for modeling and manipulating complex mathematical expressions (Chaudhuri & Cheng, 2020; Joshi et al., 2022).

THE PROSPECTS OF AI IN SOLVING COMPLEX MATHEMATICAL PROBLEMS

Artificial Intelligence (AI) is rapidly transforming the mathematical landscape, especially in tackling problems once considered unsolvable or extremely difficult. Recent advancements have demonstrated AI's capability not just as a computational tool but as a creative agent capable of proposing novel conjectures, constructing formal proofs, and exploring vast combinatorial spaces. DeepMind's AlphaTensor, for example, extends this trend by learning algorithms for matrix multiplication—a cornerstone operation in numerical mathematics—underscoring the capacity of AI to generate new mathematical knowledge (Fawzi et al., 2022). This reveals a future where AI might actively collaborate with mathematicians, augmenting both intuition and rigor.

AI models excel in areas requiring high-dimensional reasoning, such as algebraic geometry, number theory, and combinatorics. Their utility lies in pattern recognition and symbolic manipulation, which are essential in these domains. For instance, the work of Davies et al. (2021) introduced a neural-symbolic system capable of discovering and proving theorems in knot theory and group theory, previously the domain of pure mathematicians. By integrating machine learning with formal reasoning tools like Lean and Coq, AI opens new pathways for automating the theorem-proving process, making the mathematical discovery process more efficient and accessible.

Moreover, AI plays a crucial role in optimizing complex simulations and numerical methods used in applied mathematics. Neural networks and genetic algorithms are increasingly used to solve partial differential equations (PDEs), a class of problems with widespread applications in physics and engineering. Raissi, Perdikaris, and Karniadakis (2020) introduced physics-informed neural networks (PINNs), a method blending data-driven and physics-based approaches to solve nonlinear PDEs. The elegance of such models lies in their capacity to generalize, learn physical laws from data, and solve inverse problems—showing promise in areas such as fluid mechanics and quantum physics.

Despite these advances, significant challenges remain in ensuring the interpretability, reliability, and formal verifiability of AI-generated results. Mathematics, being a deductive science, demands proofs that are not just likely but logically airtight. Efforts like "formal verification" and "proof assistants" aim to bridge this gap. Bansal et al. (2021) propose architectures that combine neural networks with symbolic reasoning modules to ensure that generated outputs can be translated into formally verifiable constructs. This hybrid paradigm might be critical to integrating AI tools within rigorous mathematical workflows without compromising on precision.

Looking forward, the synergy between human intuition and AI computation promises to redefine mathematical practice. The dream of automating mathematical creativity, once considered far-fetched, is increasingly within reach. Notably, AI is being employed not just to solve problems but to pose them. For example, large language models like GPT-4 have been found capable of generating plausible conjectures and suggesting possible paths to proof, especially in fields like discrete mathematics and number theory (Polu & Sutskever, 2020). While far from perfect, this capacity hints at a collaborative model of future research where machines co-author insights with humans.

STRATEGIES OF USING AI IN SOLVING MATHEMATICAL PROBLEMS

- **Symbolic Reasoning with Neural-Symbolic Systems**

One prominent strategy involves the integration of symbolic reasoning with neural networks, often termed neural-symbolic systems. These models combine the logical rigor of symbolic math with the pattern recognition prowess of deep learning, enabling systems to not only approximate solutions but understand and manipulate algebraic and logical structures. This is especially useful for symbolic integration, solving equations, and logic-based theorem proving. For instance, Wu et al. (2021) proposed a hybrid architecture that learns symbolic manipulation rules by training a transformer model to emulate algebraic steps, showing promising results in symbolic equation solving.

- **Automated Theorem Proving via Reinforcement Learning**

Another strategy is the use of reinforcement learning (RL) in automated theorem proving. RL agents are trained to select sequences of mathematical operations that lead to proofs, akin to navigating a proof tree. This approach transforms mathematical reasoning into a search problem where the AI iteratively learns which steps are most promising. Alemi et al. (2021) showcased a system that outperformed traditional provers in terms of proof discovery efficiency by using deep RL to guide logical inference in mathematical libraries.

- **Sequence Modeling for Step-by-Step Problem Solving**

AI models like transformers and recurrent neural networks (RNNs) have been fine-tuned to generate step-by-step solutions to mathematical problems, similar to how humans solve equations line-by-line. These models are particularly adept at arithmetic word problems and algebraic manipulations. By training on large datasets of solved problems, they learn implicit rules of transformation.

- **Formal Verification with AI-Enhanced Proof Assistants**

In formal mathematics, proof assistants such as Coq or Lean have been enhanced with AI to automate parts of the proof process. By integrating natural language processing and machine learning, these assistants can predict the next lemma or theorem to apply. This hybrid approach reduces manual effort and increases proof generation efficiency in complex mathematical domains.

- **Data-Driven Mathematical Discovery**

AI is also employed in mathematical conjecture generation and pattern discovery, where deep learning models analyze large datasets of mathematical objects to identify hidden patterns or relationships. This strategy aids in hypothesizing new theorems or simplifications. Davies et al. (2021) introduced a system where AI suggested mathematical insights in knot theory and representation theory that were later verified by human mathematicians, marking a milestone in AI-assisted discovery.

THE CHALLENGES OF USING AI IN SOLVING MATHEMATICAL PROBLEM

The integration of Artificial Intelligence (AI) into mathematical problem-solving has seen significant developments, yet it continues to face various challenges. These challenges span from limitations in symbolic reasoning to difficulties in interpreting natural language in mathematical contexts. Below is a breakdown of these issues,

- **Symbolic Reasoning and Abstract Logic**

One of the core challenges for AI in mathematics is symbolic reasoning. Unlike numerical computation, solving mathematical proofs or puzzles often requires abstract reasoning, logical deductions, and symbolic manipulation that AI models typically struggle with. For instance, AI systems often fail in tasks that involve generalization beyond pattern recognition. This is emphasized by Faldu et al. (2021) who discussed the need for AI to adopt more tractable strategies to handle math word problems that require symbolic interpretation and deeper reasoning.

- **Linguistic Ambiguity in Word Problems**

Mathematical word problems are inherently tied to natural language, which introduces ambiguity that AI struggles to resolve. AI models like ChatGPT can misinterpret sentence structures, context, or mathematical intent in word problems. Daher & Gierdien (2024) showed that even with well-defined equations, ChatGPT generated incorrect answers due to misreading problem context or linguistic cues, highlighting a gap in semantic understanding.

❈ **Lack of Explainability and Transparency**

Unlike traditional mathematical proofs where reasoning can be followed step by step, AI-generated solutions often lack explainability. This raises trust and usability issues, particularly in education and research. Generative models offer solutions without justifications, thereby limiting their utility in learning environments.

❈ **Educational Limitations and Misconceptions**

From an educational perspective, the use of AI in solving mathematical problems might lead students to bypass cognitive engagement. This issue was highlighted by Elizondo-García et al. (2025), who argued that while ChatGPT can assist in challenge-based learning, it may also foster dependency if critical thinking isn't encouraged alongside.

❈ **Challenges in Problem Generalization and Heuristics**

AI models often struggle with generalizing problem-solving strategies across various domains of mathematics. They are good at recognizing previously seen formats but weak at adapting heuristics for novel structures. Davis (2024) reflects on this limitation in common sense reasoning, stressing that despite decades of development, AI still lacks robust mathematical comprehension.

Strategic Steps of Mitigating the Challenges of Using AI in Solving Mathematical Problem

Mitigating the challenges associated with using Artificial Intelligence (AI) in solving mathematical problems requires deliberate strategies, ranging from ethical frameworks to pedagogical reform. Recent research has outlined various strategic steps educators, developers, and policymakers can take.

- **Embedding Human Oversight and Ethical Frameworks**

One of the core strategies is embedding ethical principles and human intervention within AI systems. Since AI lacks contextual understanding and judgment, especially in abstract domains like mathematics, ensuring oversight can mitigate the risks of misinterpretation and error propagation. For example, Rane (2023) emphasized the importance of designing AI systems that incorporate human-in-the-loop models to increase reliability and ethical responsibility in solving mathematical problems.

- **Promoting Digital and Pedagogical Literacy among Educators**

Teachers must be adequately trained in AI-enhanced learning environments to mitigate cognitive overload and reinforce accurate mathematical thinking. Egara & Mosimege (2024) highlighted that when teachers understand AI systems, they can more effectively integrate tools like ChatGPT, helping students avoid misuse or dependency.

- **Re-thinking Data Strategies and Fairness Models**

Another approach involves rethinking how data is curated and integrated into AI systems to ensure fairness and reduce biases. Aldoseri et al. (2023) proposed robust data strategies that address bias at both the input and model stages, thereby making AI outputs more equitable and generalizable, especially in solving varied mathematical scenarios.

- **Enhancing Cognitive Scaffolding through AI Coding Tools**

Integrating AI coding into mathematics curriculum improves students' computational thinking and problem-solving skills, which supports mitigation by reinforcing understanding rather than bypassing it. The synergy between coding and math can help transform abstract reasoning into structured logical steps, thereby improving learner outcomes when AI is used.

CONCLUSION

AI has shown remarkable potency in tackling complex mathematical problems with accuracy and speed. By leveraging deep learning, symbolic reasoning, and reinforcement learning, it now contributes to solving equations, generating proofs, and discovering new patterns. These strategies have opened new avenues in both pure and applied mathematics. The prospects of AI point toward deeper collaboration with human mathematicians. As AI tools evolve, they promise to enhance creativity, not just computation. This synergy is reshaping mathematical research and problem-solving. Thus, AI stands as a powerful force driving the future of mathematics forward.

RECOMMENDATIONS

- Educational institutions should incorporate AI tools and concepts into mathematics programs to equip students with the skills to use AI for solving complex problems.
- Governments and academic bodies should fund research into hybrid AI models that combine symbolic logic with deep learning to advance AI's reasoning capabilities in mathematics.
- Educational institutions should incorporate AI tools and concepts into mathematics programs to equip students with the skills to use AI for solving complex problems.

REFERENCES

Aldoseri A., Al-Khalifa K. & Hamouda A. (2023). Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges. *Applied Sciences,* 13, 7082.

Alemi A., Irving G., Szegedy C. & Urban J. (2021). DeepMath: Deep Sequence Models for Premise Selection. *Journal of Automated Reasoning,* 65(2), 213–234.

Alemi, A. A. (2021). Learning Theorem Proving with Reinforcement Learning and Graph Neural Networks. *Journal of Automated Reasoning*. https://doi.org/10.1007/s10817-020-09575-0

Bansal, Y. (2021). Learning to reason in large theories without imitation. NeurIPS. https://arxiv.org/abs/2112.08193

Chaudhuri, S., & Cheng, K. (2020). Neural-symbolic systems for mathematical reasoning. *Artificial Intelligence Review*, 53, 4199–4226. https://doi.org/10.1007/s10462-020-09834-6

Choi, J. (2021). Neuro-symbolic approaches for reasoning in complex domains. *Nature Machine Intelligence*, 3(7), 593–602. https://doi.org/10.1038/s42256-021-00336-9

Cunningham, D., et al. (2023). Symbolic Discovery of Optimization Algorithms. arXiv. https://arxiv.org/abs/2303.17581

Daher W. & Gierdien F. (2024). Use of Language By generative AI Tools in Mathematical Problem Solving: The Case of ChatGPT. *African Journal of Research in Mathematics, Science and Technology Education*, 2024. https://scholar.sun.ac.za/server/api/core/bitstreams/281ed820-27b9-433c-bb12-3b37948ddaba/content

Davies A., Bosma J. & Mehta A. (2021). Advancing Mathematics by Guiding Human Intuition with AI. *Nature,* 600(7887), 70–74.

Davies, A. (2021). Advancing mathematics by guiding human intuition with AI. *Nature*, 600, 70–74. https://doi.org/10.1038/s41586-021-04086-x

Egara F. & Mosimege M. (2024). Exploring the Integration of Artificial Intelligence-Based ChatGPT into Mathematics Instruction: Perceptions, Challenges, and Implications for Educators. *Education Sciences,* 14(7): 742.

Elizondo-Garcia M., Hernandez-De la Cerda H., Benavides-Garcia I., Caratozzolo P. & Membrillo-Hernandez J. (2025). Who is solving the challenge? The use of ChatGPT in mathematics and biology courses using challenge-based learning. *Front. Edu.*, 10, 2025.

Davis E. (2023). Mathematics, word problems, common sense, and artificial intelligence. Available at: https://arxiv.org/pdf/2301.09723

Faldu K., Sheth A., Kikani P., Gaur M. & Avasthi A. (2021). Towards Tractable Mathematical Reasoning: Challenges, Strategies, and Opportunities for Solving Math Word Problems. Available at: https://arxiv.org/pdf/2111.05364

Fawzi, A. (2022). Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature*, 610(7930), 47–53. https://doi.org/10.1038/s41586-022-05172-4

Joshi, S., et al. (2022). Mathematical Reasoning via Self-supervised Skip-tree Networks. arXiv. https://arxiv.org/abs/2203.15709

Kaitera, S. and Harmoinen, S. (2022) Developing mathematical problem-solving skills in primary school by using visual representations on heuristics. *LUMAT Special Issue*, 10(2), 111–146

Lample, G., & Charton, F. (2020). Deep learning for symbolic mathematics. *Nature*, 588(7836), 66–70. https://doi.org/10.1038/s41586-020-2913-0

Polu, S., & Sutskever, I. (2020). Generative Language Modeling for Automated Theorem Proving. https://arxiv.org/abs/2009.03393

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2020). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

Rane N. (2023). Enhancing Mathematical Capabilities through ChatGPT and Similar Generative Artificial Intelligence: Roles and Challenges in Solving Mathematical Problems. Available at: https://ssrn.com/abstract=4603237

Staff. C. (2024) What Is Artificial Intelligence? Definition, Uses, and Types. https://www.coursera.org/articles/what-is-artificial-intelligence

Stryker, C. (2024) What is artificial intelligence (AI)? https://www.ibm.com/think/topics/artificial-intelligence#:~:text=Eda%20Kavlakoglu,a%20self%2Ddriving%20car).

Tian, Y., et al. (2022). Formal Mathematics Statement Curriculum Learning. arXiv. https://arxiv.org/abs/2203.13379

Vye, N. J., Goldman, S. R., Voss, J. F., Hmelo, C., and Williams, S. (2024). Cognition and Technology Group at Vanderbilt. "Complex Mathematical Problem Solving by Individuals and Dyads." Cognition and Instruction 15(4), 435–84. http://www.jstor.org/stable/3233775.

Wu Z., Jin Y. & Zhu J. (2021). Neural-Symbolic Machines for Algebraic Reasoning. *Proceedings of the AAAI Conference on Artificial Intelligence,* 35(9), 7902–7910.

Zheng, D., et al. (2021). minif2f: A benchmark for formal mathematical reasoning. arXiv. https://arxiv.org/abs/2109.07740