nature reviews physics
https://doi.org/10.1038/s42254-024-00740-1

Perspective

Check for updates

AI-driven research in pure mathematics and theoretical physics

Yang-Hui He[1,2]

Abstract

The past five years have seen a dramatic increase in the usage of artificial intelligence (AI) algorithms in pure mathematics and theoretical sciences. This might appear counter-intuitive as mathematical sciences require rigorous definitions, derivations and proofs, in contrast to the experimental sciences, which rely on the modelling of data with error bars. In this Perspective, we categorize the approaches to mathematical and theoretical discovery as 'top-down', 'bottom-up' and 'meta-mathematics'. We review the progress over the past few years, comparing and contrasting both the advances and the shortcomings of each approach. We believe that although the theorist is not in danger of being replaced by AI systems in the near future, the combination of human expertise and AI algorithms will become an integral part of theoretical research.

Sections

[1]London Institute for Mathematical Sciences, Royal Institution, London, UK. [2]Merton College, University of Oxford, Oxford, UK. e-mail: hey@maths.ox.ac.uk

Introduction

The past five years have seen a lot of progress in the use of artificial intelligence (AI) in theoretical investigations in pure mathematics and theoretical physics. This domain is markedly different from how AI technology has been transforming a wide array of human activity in the past decade[1–3]. For practical purposes, many human tasks only require black-boxes: learning from trial-and-error and performing within a margin of error. For example, the primary goal of a medical doctor is to cure the patient; understanding the exact mechanisms of the disease is secondary. Black-box machine learning based on large statistical samples is exactly what deep neural networks do best. However, this approach is unsatisfactory for the scientist, whose job is to understand and to question, and even more so for the theorist.

In this Perspective, we use the term 'theoretical science' to include both pure mathematics and the development and testing of theories and hypotheses using mathematics, exemplified by theoretical physics. In other words, we adhere to the British academic convention and consider the fields such as theoretical physics or theoretical computer science as subdisciplines of mathematics.

Although experimental physicists at CERN, searching for new particles, were among the first scientists to use AI[4] in the 1990s, it was not until 2017–2018 that machine-learning techniques emerged in theoretical physics[5–9] and in pure mathematics[10,11]. The desire for interpretability and explicability is likely the reason why the use of AI in theoretical science has been a relative latecomer compared with other areas of science.

But this field is catching up fast. In theoretical physics, examples of AI-driven research include particle phenomenology from string theory[12–20], establishing dictionaries between field theory and deep learning[21–23], theoretical cosmology[9,24,25], quantum field theories[26–30] and uncovering fundamental symmetries[31–36]. In parallel, in pure mathematics, examples span algebraic geometry[5,10,37–44], algebraic structures and representation theory[45–47], symbolic algebra and computation[48–54], differential and metric geometry[55–57], number theory[58–61], graph theory and combinatorics[43,62,63], to knot theory[47,64,65].

In general, the question of what role AI technology will play in society is approached with a mixture of optimism[66–73] and anxiety[2,3,74,75]. The purpose of this Perspective is to overview and discuss the advances that have been made in the past few years in the use of AI in theoretical science[10,71,76–80], highlighting the concerns and limitations[81–84]. We caution against both overenthusiasm and unwarranted fear and convey that the future of the mathematical sciences will be an inevitable, and hopefully harmonious, union between human and AI.

AI-driven mathematics and theoretical science

To answer the question of what is AI-driven or AI-assisted theoretical research, one needs to first ask a more basic question: how do mathematicians do mathematics? Without delving into the philosophy of mathematics or the philosophy of science, we take a pragmatic approach to problem-solving. In an earlier review[85], we discussed two opposite directions of how machine learning could help with uncovering mathematical structures. In the light of the advances in large language models (LLMs), it is expedient to add a third direction. In some ways, this trio echoes the three major philosophical schools[86] of mathematics (formalism, logicism and intuitionism) at the turn of the last century, to which we pay tribute. The reader is also referred to a recent information-theoretic treatment[73] of what a mathematician is. We classify the approaches to doing mathematics/theoretical science as follows:

- Bottom-up: one can think of mathematics as being built from foundational axioms, in which all theorems and equations are constructed from the roots up using logic. We refer to this approach as 'bottom-up' to reflect the rigorous nature of theoretical research. This is somewhat in the spirit of Hilbert's Formalism Programme[87] (Box 1) and the logicism of Russell–Whitehead[88], in which mathematics is formulated by acting precise rules on well-defined symbols.
- Meta-mathematics: meta-science has been used to refer to the science of scientific research[89]. In the context of this Perspective, it refers to looking at mathematics from a distance, especially from the corpora of published papers — without necessarily comprehending the actual details. (The term is also different from its usage by John McCarthy in the context of mathematical logic in AI[90].)
- Top-down: the practicing theorist often 'experiments' and 'conjectures' before tackling a proof or derivation. We will refer to this as 'top-down', in which an overarching view, based on experience and speculation, guides one towards a problem in which truth can only be attained in the human mind. (This approach, though somewhat in the spirit of, should be distinguished from Brouwer's intuitionist mathematics[91].) We will argue that even a black-box AI can help make serious progress in formulating precise conjectures, in which rigour arises from statistical inference[71].

There is an important point on terminology. In philosophy, top-down means less theoretical to more theoretical. Here, we are borrowing the words top-down and bottom-up from physics, where the starting point of bottom-up model building is from low-energy scale and that of top-down is from high-energy scale.

In the next sections, we will discuss in detail how AI has been instrumental to each of these approaches.

Bottom-up mathematics

Hilbert's programme[87] of building up mathematical truths from the ground formulated in the early 1920s received a fatal blow a decade later from the undecidability and incompleteness theorems of Gödel[92], Church[93] and Turing[94]. These essentially showed that within any mathematical system, there are statements whose truth value cannot be decided. Nevertheless, the space of decidable and interesting statements is so vast that one can certainly focus first on these. Thus, despite the logical impossibility of building all mathematical statements bottom-up, theorists never stopped pursuing proofs for countless propositions. This led to the modern-day answer to Alfred North Whitehead's and Bertrand Russell's work on the foundations of mathematics 'Principia Mathematica'[88]: the automated theorem proving programme (ATP). Arguably, the first AI system for mathematics — or indeed, the first AI system — was the logic theory machine[95], an early computer system created by Newell, Simon and Shaw in 1956, around the same time as the emergence of the first trainable neural network[96]. The logic theory machine succeeded in proving a number of propositions of the Principia Mathematica.

Over the second half of the twentieth century, it became clear that an increasing number of proofs of fundamental results in mathematics are impossible without the computer. These have ranged from situations in which key steps reduce to extensive brute-force computation, such as in the four-colour theorem, to more extreme circumstances in which it takes longer than a human lifespan to go through all the details, such as the classification of simple finite groups. The dependence of the human theorist on machines has prompted such influential

Box 1 | Mathematics glossary

Four colour theorem: one of the most famous standing problems in mathematics conjectured in the 1850s and only settled in 1976 by Kenneth Appel and Wolfgang Haken. It states that any map (in a plane) can be coloured with at most four different colours such that no neighbouring countries have the same colour. This can be phrased in terms of graphs theory as: the vertices of every planar graph can be labelled with at most four labels such that no two adjacent vertices receive the same label.

Polynomial Freiman–Ruzsa conjecture: in the field of additive combinatorics, a central problem is concerned with the so-called doubling constant K: given a set A of integers, $A+A$ is the set of all possible pairwise sums, and K is the ratio of the size of $A+A$ to that of A. Freiman's theorem roughly states that if K is small, then the elements of A, when sorted, grow linearly. The polynomial Freiman–Ruzsa conjecture, formulated by Katalin Marton, is a generalization of Freiman to abelian groups.

Fermat's last theorem: perhaps the most celebrated statement in mathematics, Fermat's last theorem was a proposition made without proof formulated by Pierre de Fermat in the seventeenth century: no three positive integers a, b and c satisfy the equation $a^n+b^n=c^n$ for any integer $n>2$. (The case of $n=1$ is trivial with infinitely many solutions, and $n=2$ also has infinitely many solutions that constitute integer sides of Pythagoras' right-angle triangle, such as $(a, b, c)=(3, 4, 5)$.) The final proof, made by Andrew Wiles in 1994, had to use a host of sophisticated results from modern mathematics.

Birch–Swinnerton-Dyer conjecture: a Millennium Prize problem, the conjecture of Birch–Swinnerton-Dyer (BSD), is concerned with integer points on an elliptic curve, a cubic polynomial in two variables, such as $y^2=x^3+1$. Elliptic curves are central to modern mathematics. For example, Wiles' proof of Fermat's last theorem relied on it, being phrased in terms of integer points on a particular elliptic curve (called the Frey curve). BSD conjectures that the number (called the rank) of infinite families of integer points on any elliptic curve, for example, $y^2=x^3+1$, can be determined by the analytic properties of a function (called the L-function), which encodes the solutions of the curve when reduced over each prime p. Here, using the aforementioned example, reduction over p means to find integers (x, y) such that $y^2 \equiv x^3+1 \pmod{p}$.

Murmurations conjectures: relevant to BSD, the recently formulated murmuration conjectures constitute a first non-trivial discovery in number theory based on human–artificial intelligence collaboration. It states that L-functions for elliptic curves defined earlier, averaged over different curves in a certain range, show an oscillatory (murmuration) behaviour. Curves of different ranks have different types of murmurations, reflecting BSD. This murmuration is expected to hold also for other classes of L-functions.

Riemann hypothesis: the Riemann hypothesis is another Millennium Prize problem and the most important unsolved mathematical problem. Tens of thousands of theorems, including many on such fundamental results as the distribution of primes, rely on it, being true. Proposed in 1859 by Bernhard Riemann, there have been hundreds of equivalent formulations thereof. Here is perhaps the simplest one. Consider the Möbius mu-function alluded to in the text. For a positive integer n, it is defined as follows: $\mu(n)=0$ if n has any repeated factors in its prime factorization (for example, $\mu(12)=\mu(2^2 \cdot 3)=0$ because 2 is repeated); otherwise, it is (-1) to the power of the number of distinct prime factors (for example, $\mu(42)=\mu(2 \cdot 3 \cdot 7)=(-1)^3=-1$). In other words, μ detects the parity of distinct prime factors of n. Now consider the function $M(x)=\sum_{n \le x} \mu(n)$. The Riemann hypothesis states that $M(x)$ roughly grows as $\sqrt{x}$. (More strictly, that for every $\varepsilon>0$, $|M(x)| \le Kx^{\frac{1}{2}+\varepsilon}$ for some positive real number K as $x \to \infty$.)

Hilbert's programme: to avoid paradoxes and inconsistencies which have arisen in the preceding century, David Hilbert in the 1920s proposed to ground all existing mathematics to a finite, complete set of axioms and provide a proof that these axioms be consistent. This programme was shown to be essentially unachievable by the incompleteness theorems of Gödel in the 1930s.

Brouwer's intuitionist mathematics: in the philosophy of mathematics, intuitionism is an approach in which mathematics is considered to be purely the result of constructive, human mental activity, rather than the discovery of fundamental principles that exist in an objective reality. The mathematician and philosopher Luitzen Brouwer proposed in the 1920s of the 'subjectivity' of mathematics that a mathematical statement corresponds to a mental construction, and a mathematician can assert the truth of a statement only by verifying the validity of that construction by intuition.

figures as Terence Tao[97], and addresses at the International Congress of Mathematicians[98], to seriously consider the future of mathematics.

Although the first proof-assistant[99] appeared in the 1970s, Isabelle/HOL[100], Coq[101], Agda[102] and Lean[103] softwares are spear-heading the ATP programme in this century. One notable direction well underway is the Xena project (https://xenaproject.wordpress.com/) aiming to formalize all (every statement and every step of proof) of undergraduate-level mathematics into Lean. Last year, Lean's MathLib library was used to prove the polynomial Freiman–Ruzsa conjecture[104]. The effort to establish a full database of all of the contemporary mathematics in Coq or Lean format is still far from completion (private communications with Kevin Buzzard and James Davenport). For example, only earlier this year was a new project launched to formalize all the requisite pieces to Wiles' proof of Fermat's last theorem. This suggests that having AI automation on selecting correct proof strategies given a proposition or conjecture is still a way to come (private communication with Kevin Buzzard and James Davenport). Nevertheless, such a mathematics database will enable data mining for new theories, which brings us to the next point.

Meta-mathematics

From the Principia Mathematica to the advancement of computer science, there has been a tradition of viewing mathematics as a language[105].

Inspired by Ludwig Wittgenstein[106], one can think of 'Mathematics as Language' and any proposition is just a set of symbols, led to by sequences of symbols that one calls proof or derivation.

Indeed, the field of Natural Language Processing (NLP) is rooted in Turing's original proposal of his eponymous test[107]. Advances in AI technology and the availability of internet data have propelled NLP to the era of the LLMs. Chatbots built on LLMs have passed the Turing test[108]. The important point here is that LLMs have no 'understanding' of the underlying material; these models merely group together words in the right order based on large corpora of statistical samples. The philosophical meaning of 'understanding' aside (and this is why we chose to call this direction 'meta-mathematics'), it is clear that LLMs have been transformative in mimicking human communication.

One of the earliest experiments[109] using an LLM for theory was the application of the Word2Vec[110] neural network (perhaps the most basic LLM technique) to the titles of several sections of the arXiv preprint server, the most comprehensive repository for contemporary research in mathematics and theoretical physics. Perhaps more interesting than in retrieving seemingly sensible linguistic identities (for example, 'string-theory + Calabi–Yau = M-theory + G2') was a comparison with the viXra preprint server, the repository of fringe ideas not accepted by main-stream science. From the titles alone, one could significantly distinguish (from the confusion matrix) different subfields of theoretical physics (high-energy theory, high-energy phenomenology, general relativity/quantum cosmology and so on) in arXiv, whereas in viXra this is not the case. In other words, the syntax of proper theoretical science is more self-consistent than that of fringe science even at the level of titles. (By enriching the data with the inclusion of abstracts, the application of Word2Vec on papers in material science uncovered new chemical reactions[111].)

Today, Word2Vec has been superseded by other deep learning architectures such as transformers, and the programme of LLM for mathematics is blooming[112–119]. Notably, in parallel to projects led by OpenAI[120] and MetaAI[121] in the space of mathematics education, DeepMind's AlphaGeo[116] has recently been able to generate correct, human-understandable proofs for Olympiad-level problems in Euclidean geometry. However, one should note that a non-trivial part of the success of AlphaGeo involved old-fashioned search using the AR+DD (Algebraic Rules and Deductive Database). These advances suggest that when and if a complete linguistic database in the precise format of, say, 'Lean' will exist for all contemporary mathematics, then the LLM approach to this vast data set should produce new mathematics.

Top-down mathematics

Everything we have discussed so far has to do with building correct mathematical statements. But frequently, one has no idea what statement one should try to demonstrate. Indeed, how does the practicing mathematician actually work? In many ways, our papers are written backwards. From day to day, we doodle on paper and on board, experimenting with ideas, mistakes and expressions, until something sensible comes out. Then, we go back and formalize with definitions followed by theorems and derivations that lead to logical conclusions. Thus, journal papers in mathematics and theoretical physics look bottom-up, even though the discovery process is quite the opposite. The duality between these two directions is called 'synthesis versus analysis'[122]. Historically, this is how some of the greatest theoretical discoveries were made. Isaac Newton and Leonhard Euler were freely manipulating formal expressions in calculus, centuries before a proper notion of analysis and convergence; Evariste Galois showed the unsolvability of the quintic by radicals by seeing the structures of permutation groups, before the definitions of groups and fields taught today. In physics, theorists freely manipulate Feynman integrals to obtain results that agree with experiment to astounding accuracy, even though a mathematically rigorous formulation of quantum field theory is lacking.

Over the past couple of years, the worry that AI will replace the human mathematician has been growing[123]. Some see the human mathematician as a bottom-up Logical Theory Machine, building sentences (proofs and derivations) from definitions. In reality, actual mathematical research is based on a combination of inspiration, intuition and experience. In contrast with the dry bottom-up narrative, we call this almost fuzzy, aesthetics-driven approach[124] top-down. Of course, at the end of the day, all statements must be rigorous and any fuzziness and inaccuracies must be distilled out (see a recent Perspective[71]). With this in mind, let us discuss this approach in detail, with illustrative examples.

Perhaps contrary to common conception, an indispensable component to even the purest mathematical discovery is data. This is not experimental data in the sense of, say, particle trajectories from CERN, with errors and variance, but results of classifications and computations, for example, tables of characters of finite groups. These 'pure data' are exact and without statistical error and shed light on the underlying theory. To quote mathematician Vladimir Arnold[125], 'mathematics is the part of physics where experiments are cheap'.

Confronted with the ancient problem of finding patterns in primes, which dates back at least to the time of Euclid, 16-year-old Carl Friedrich Gauss defined the prime counting function, which gives the number of prime numbers not exceeding a positive real $x \in \mathbb{R}_+$

$$\pi(x) = \#\{p \le x : p \ \text{ prime}\}. \tag{1}$$

He consulted tables available at the time and computed tens of thousands more (by hand!) and simply plotted $\pi(x)$. Gauss conjectured that $\pi(x) \sim x/\ln(x)$. This profound observation had to wait for the establishment of complex analysis by Augustin-Louis Cauchy and Bernhard Riemann to be proven by Jacques Hadamard and Charles Jean de la Vallée Poussin in 1896. It is now known as the prime number theorem, one of the most important results in mathematics.

In the twentieth century, mathematicians Bryan John Birch and Peter Swinnerton-Dyer plotted, using the early computers of the 1960s, ranks and other quantities for elliptic curves and conjectured that the order of vanishing of the L-function $L(s)$ for the curve (which encodes the number of integer points on it modulo different prime numbers) at $s \to 1$ equals to the rank. This observation is known as the Birch–Swinnerton-Dyer conjecture and is a Millennium Prize problem[126] central to modern mathematics.

The above are but two of the countless examples in which experimenting with pure data can lead to profound results. They illustrate the importance of conjectures. In theoretical research, finding the interesting problem is vital, and this process is often guided by the almost undefinable process of intuition. Godfrey Harold Hardy's definition of mathematics[127] is succinct: 'A mathematician, like a painter or a poet, is a maker of patterns'. Most scientists who study real world data from observations would first distil the problem into a mathematical setting. Then it again becomes a mathematical game of pattern spotting, from graphs and plots, to formal symbols. But there is one thing that AI systems can do better than humans, that is pattern recognition, especially when the data are in high dimensions. This is exemplified in Box 2.

Box 2 | Playing with binary sequences

Let us perform the following simple experiment. First, given the sequence {0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1} and asked what the next number is, any human would instantly say 0. One way to explain why is that the sequence shows whether n divides 3, for positive integers $n \in \mathbb{Z}_{>0}$. Next, try {0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0}. An inspired person might, after some experimentation, conclude that the next number is 1; this is the sequence of PrimeQ, whether the nth positive integer is prime or not. Finally, try {1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1}. This is the sequence of whether the nth positive integer has an even (0) or odd (1) number of prime factors counted with multiplicity, a shifted version of the so-called Liouville lambda-function (see a recent artificial intelligence (AI) treatment[144] of the closely related Möbius mu-function). Uncovering patterns in the lambda-function would have incredible repercussions for mathematics: there are equivalent formulations of the Riemann hypothesis in terms of this sequence[145].

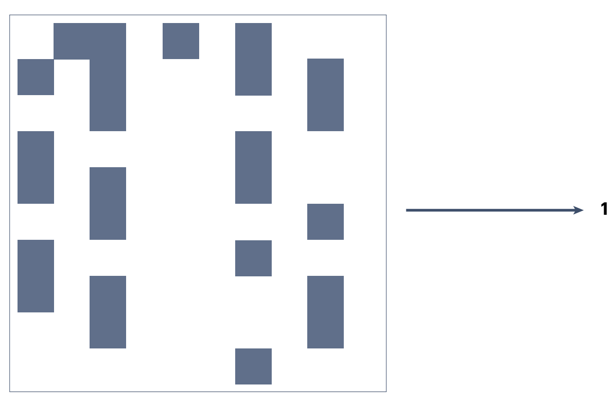
What if one gave the sequence to an AI system to solve (for instance, using a supervised machine learning (ML) algorithm)? To establish a reasonable training set, one could choose the following representation (and indeed the choice of representation is extremely important). Take one of the aforementioned infinite sequences $\{a_i\}_{i=1,2,3,...}$ and a sliding window of length N. In other words, consider a set of sequences $\{\{a_i\}_{i=1,2,...,N}, \{a_i\}_{i=2,3,...,N+1}, \ldots \{a_i\}_{i=k,k+1,...,N+k-1}\}$ for some k. Here, k will be taken to be sufficiently large (say 100,000) to create a decent-size data set, and N will be taken to be sufficiently large (say 100) to give enough features. (After all, mathematical data are cheap.) One can then consider each of the finite subsequences as a single vector in $\mathbb{R}^N$ and label it by the next number outside the window:

$$\begin{aligned}\{&(a_i)_{i=1,2,...,N} \longrightarrow a_{N+1},\\ &(a_i)_{i=2,3,...,N+1} \longrightarrow a_{N+2}, \quad \ldots,\\ &(a_i)_{i=k,k+1,...,N+k-1} \longrightarrow a_{N+k}\}.\end{aligned} \quad (2)$$

Note that one has chosen the sequences judiciously to standardize everything into a binary classification problem of binary vectors of dimension N. The question is then: having seen k labelled samples, how well will the ML algorithm predict on k' unseen vectors? This familiar supervised ML paradigm can then be compared with the human performance on the same task.

One can readily check that with the most basic ML algorithms suited for this problem, such as decision trees, support vector machines, relatively shallow feed-forward neural networks with rectified linear unit activation functions and so on, on equation (2) applied to the three sequences. For the first one, 100% accuracy is reached very quickly for any of the ML algorithms. For the second, one reaches about 80% (in this case because of the increasing rarity of primes — approximately by a factor of $x/\ln(x)$ owing to the prime number theorem — window size needs to be scaled accordingly), and for the third one struggles to find any AI algorithm that would beat 50%. What this means is that the first sequence, a trivial problem for humans, is as trivial for the AI system; for the second, the AI approach might be finding some version of the Sieve of Eratosthenes for checking PrimeQ; and for the last the AI system is not beating a random guess. Of course, should one find an algorithm which does, then one might be well underway in finding a new approach towards the Riemann hypothesis.

Many of the papers referenced in the introduction use similar ideas to the previous paragraph, but to much more sophisticated situations. Indeed, what (computable) mathematics, in the sense of a Turing machine, does not fall into some version of equation (2)? We can make the situation even more visual and suited to AI by wrapping each N vector into an $m \times n = N$ matrix, which can be interpreted as a pixelated image in which 1 is black and 0 is white. For instance, suppose $N = 100$, the first vector for the second sequence can be wrapped into a 10×10 matrix, together with a label 1 (since 101 is prime), illustrated subsequently. Take, as a much more elevated example, the problem of computing a topological invariant for a manifold in algebraic geometry (which involves advanced calculations). Yet, one can represent the manifold as a pixelated image by tensorizing the multidegree information of the manifold as an algebraic variety[5,10]. In a similar manner, one could reframe any mathematical computation as an image recognition problem. Learning and gaining experience and intuition from many calculations — as mathematicians and theorists do during their careers — is in analogy to training a neural network. We could summarize this paradigm loosely as: bottom-up mathematics (and meta-mathematics) is language processing, whereas top-down mathematics is image processing.



The Birch test

The key steps to top-down mathematics are first, identifying the problem and second, identifying a strategy to attack the problem. Both depend on experience, with a healthy dose of intuition. Although LLMs applied to databases such as Lean's MathLib[103] are making baby-steps in the latter, the former is formally known as conjecture formulation, exemplified by the aforementioned cases of Gauss and Birch–Swinnerton-Dyer. Can AI systems assist in telling a good conjecture from a useless one? Which patterns found from mathematical data lead to interesting, as opposed to, trivial mathematics? AI-guided conjecture formulation has been given much systematic thought[47,85,128–133].

In a 6-month workshop in Cambridge in 2023 which I helped to co-organize, participants attempted to give some criteria on AI-driven theoretical discovery and, in particular, on AI-assisted conjectures. The following criteria for AI-assisted discovery in theoretical science were identified:

- Automaticity: the discovery is completely made by AI from pattern-spotting, without any human intervention;
- Interpretability: any statements — conjectures or conclusions — must be precise to a human mathematician, who cannot distinguish them from the ones given by a human colleague;
- Non-triviality: the insight is non-trivial enough that the community of human experts will work on it.

Because the points were inspired by a talk given by Birch[134], the set of criteria is referred to as the Birch test[135]. These are very stringent criteria and so far no AI-assisted theoretical discovery has passed all three parts of the test. We now highlight with some examples in which they succeeded and failed.

Take the early experiments of obtaining topological invariants by deep-learning algebraic varieties[5,10]. They have been improved to >99.9% accuracy[41], which hints at underlying and yet-unknown structures in algebraic geometry that facilitate calculations without recourse to standard and computationally expensive methods in long-exact sequences and Gröbner bases[136]. These results suffer from the typical shortcoming of deep neural network approaches: there is no interpretable formula one could extract. Thus, they fail the interpretability condition of the Birch test. A better situation[45] is where a support vector machine found a separation between simple and non-simple finite groups by plotting the Cayley multiplication tables. However, the hypersurface of separation is so complicated and deforms further with the addition of more samples of groups that the interpretability condition is still not fulfilled. The knot invariant relations found by saliency analyses[47] and the Reidemeister moves needed to untangle extremely complicated knots[64], although novel, interesting and precise, were either already proven or have not become sufficiently influential in the field; thus they fail the non-triviality condition. Likewise, the continued fraction identities found by the Ramanujan machine[130] or the physical conservation laws found by AI Feynman[33] also belong to this category. Even in a practical task, the faster matrix multiplication algorithm found by DeepMind[137] was shortly thereafter beaten by an algorithm devised by human researchers[138].

The closest any AI-guided theoretical discovery to fulfil the three criteria is the murmuration conjectures in the number theory[139]. This approach fulfils the interpretability and non-triviality but fails the automaticity because human researchers intervened in the process by digging under the hood. Surprised by why AI was doing so well at distinguishing ranks of elliptic curves in the context of the Birch–Swinnerton-Dyer conjecture, the researchers had to home in on a principal component analysis and then look at the weight matrices to extract an unexpected formula.

Outlook

The human theorist is not in danger of being put out of the job in the foreseeable future. From the lack of a complete bottom-up MathLib database[103] for all of mathematics, to the challenges LLMs would face given the vast search space of proof strategies even with such a database, to the exacting requirements of the Birch test in top-down mathematics, we are far from automating theoretical discovery.

Nevertheless, it is undeniable that AI is beginning and will continue to play a pivotal role in partnership with the human mathematician and theoretical scientist. In each of our three directions we discussed in this Perspective, there is both pressing and long-term work to be done.

- For bottom-up, after the formalization of Wiles' proof, it would be expedient to start the formalization and cross-check the classification theorem of finite simple groups, one of the foundational and yet never scrutinized theorems in mathematics (the entire proof runs over some 10,000 pages)[140]. This is likely to take a number of years as currently Lean[103] has more arithmetic geometry than other branches of mathematics.
- For meta-mathematics, the clear path ahead is a concerted effort by the scientific community to use LMMs to mine the knowledge on the arXiv. A primitive version was done for disciplines related to mathematical physics some time ago[109], and arXiv also has its internal language models that, for instance, auto-classifies submissions. Now, Llama[141] has included the arXiv in its LLM. It would be great to extract new mathematical ideas using Llama. That said, meta-mathematics is challenging. For example, one could ask ChatGTP to differentiate a function and it would do so quite accurately (without doing any mathematics). However, when faced with a seemingly much simpler question of, for example, 'find the 20th digit in the decimal expansion of 7/11', it would fail miserably and not outperform a 10% random guess. This question can only by answered by actually doing long division. Even throwing more compute to 'LLM for maths' would not solve this issue; one needs to know when to invoke genuine mathematical software, as is beginning to be done in Wolfram Alpha in conjecture with ChatGTP[142].
- For top-down mathematics, we are on the verge of passing the Birch test. In the coming years, one would imagine the fully automated generation of non-trivial conjectures. The challenge here is to have a system that selects them and have them guide human mathematicians. This hope resonates well with the comments of mathematician Jordan Ellenberg[143]: 'Some people imagine a world where computers give us all the answers [in mathematics]. I dream bigger. I want them to ask good questions'.

In the eighteenth and nineteenth centuries, Gauss's intuitions alone were good enough to spot patterns that led to such profound results as the prime number theorem. In the twentieth century, computer experimentations were needed alongside the insights of Birch and Swinnerton-Dyer to raise their conjecture. In the twenty-first century, AI systems will help human experts to find new insights, conjectures and strategies for derivations and proofs.

Published online: 05 August 2024

References

1. Sejnowski, T. J. *The Deep Learning Revolution* (MIT Press, 2018).
2. Acemoglu, D., Autor, D., Hazell, J. & Restrepo, P. Artificial intelligence and jobs: evidence from online vacancies. *J. Labor Econ.* **40**, S293–S340 (2022).
3. Bruun, E. P. & Duka, A. Artificial intelligence, jobs and the future of work: racing with the machines. *Basic Income Stud.* **13**, 20180018 (2018).
4. Perret-Gallix, D. & Wójcik, W. New computing techniques in physics research. In *Proc. 1st Int. Workshop on Software Engineering, Artificial Intelligence and Expert Systems in High-energy and Nuclear Physics* (CNRS, 1990).
5. He, Y.-H. Deep-learning the landscape. *Phys. Lett. B* **774**, 564–568 (2017).

6. Carifio, J., Halverson, J., Krioukov, D. & Nelson, B. D. Machine learning in the string landscape. *J. High Energy Phys.* **09**, 157 (2017).
7. Krefl, D. & Seong, R.-K. Machine learning of Calabi–Yau volumes. *Phys. Rev. D* **96**, 066014 (2017).
8. Ruehle, F. Evolving neural networks with genetic algorithms to study the string landscape. *J. High Energy Phys.* **08**, 038 (2017).
9. Liu, J. Artificial neural network in cosmic landscape. *J. High Energy Phys.* **12**, 149 (2017).
10. He, Y.-H. *The Calabi–Yau Landscape: From Geometry, to Physics, to Machine Learning* (Springer Cham, 2018).
11. Bull, K., He, Y.-H., Jejjala, V. & Mishra, C. Machine learning CICY threefolds. *Phys. Lett. B* **785**, 65–72 (2018).
12. Demirtas, M., Long, C., McAllister, L. & Stillman, M. The Kreuzer–Skarke axiverse. *J. High Energy Phys.* **04**, 138 (2020).
13. Mütter, A., Parr, E. & Vaudrevange, P. K. S. Deep learning in the heterotic orbifold landscape. *Nucl. Phys. B* **940**, 113–129 (2019).
14. Cole, A. & Shiu, G. Topological data analysis for the string landscape. *J. High Energy Phys.* **03**, 054 (2019).
15. Otsuka, H. & Takemoto, K. Deep learning and k-means clustering in heterotic string vacua with line bundles. *J. High Energy Phys.* **05**, 047 (2020).
16. Ruehle, F. Data science applications to string theory. *Phys. Rep.* **839**, 1–117 (2020).
17. Larfors, M. & Schneider, R. Explore and exploit with heterotic line bundle models. *Fortsch. Phys.* **68**, 2000034 (2020).
18. Bies, M. et al. Machine learning and algebraic approaches towards complete matter spectra in 4*d* F-theory. *J. High Energy Phys.* **01**, 196 (2021).
19. Perez-Martinez, R., Ramos-Sanchez, S. & Vaudrevange, P. K. S. Landscape of promising nonsupersymmetric string models. *Phys. Rev. D* **104**, 046026 (2021).
20. Abel, S., Constantin, A., Harvey, T. R. & Lukas, A. String model building, reinforcement learning and genetic algorithms. In *Nankai Symp. Mathematical Dialogues: In Celebration of S.S. Chern's 110th Anniversary* https://arxiv.org/abs/2111.07333 (2021).
21. Hashimoto, K., Sugishita, S., Tanaka, A. & Tomiya, A. Deep learning and the AdS/CFT correspondence. *Phys. Rev. D* **98**, 046019 (2018).
22. de Mello Koch, E., de Mello Koch, R. & Cheng, L. Is deep learning a renormalization group flow? *IEEE Access* **8**, 106487–106505 (2020).
23. Erbin, H., Lahoche, V. & Samary, D. O. Non-perturbative renormalization for the neural network-QFT correspondence. *Mach. Learn. Sci. Tech.* **3**, 015027 (2022).
24. Jinno, R. Machine learning for bounce calculation. Preprint at https://arxiv.org/abs/1805.12153 (2018).
25. Rudelius, T. Learning to inflate. *JCAP* **2019**, 044 (2019).
26. Chen, H.-Y., He, Y.-H., Lal, S. & Majumder, S. Machine learning lie structures & applications to physics. *Phys. Lett. B* **817**, 136297 (2021).
27. Kaspschak, B. & Meißner, U.-G. Neural network perturbation theory and its application to the Born series. *Phys. Rev. Res.* **3**, 023223 (2021).
28. Harvey, T. R. & Lukas, A. Quark mass models and reinforcement learning. *J. High Energy Phys.* **08**, 161 (2021).
29. Gupta, R., Bhattacharya, T. & Yoon, B. in *Artificial Intelligence for Science* (eds Choudhary, A., Fox, G. & Hey, T.) Ch. 26 (World Scientific, 2023).
30. Lal, S., Majumder, S. & Sobko, E. The R-mAtrIx Net. Preprint at https://arxiv.org/abs/2304.07247 (2023).
31. Gal, Y., Jejjala, V., Mayorga Peña, D. K. & Mishra, C. Baryons from mesons: a machine learning perspective. *Int. J. Mod. Phys. A* **37**, 2250031 (2022).
32. Krippendorf, S. & Syvaeri, M. Detecting symmetries with neural networks. *Mach. Learn. Sci. Technol.* **2**, 015010 (2020).
33. Udrescu, S.-M. & Tegmark, M. AI Feynman: a physics-inspired method for symbolic regression. *Sci. Adv.* **6**, eaay2631 (2020).
34. Cornelio, C. et al. Combining data and theory for derivable scientific discovery with AI-Descartes. *Nat. Commun.* **14**, 1777 (2023).
35. Liu, Z. et al. Machine-learning nonconservative dynamics for new-physics detection. *Phys. Rev. E* **104**, 055302 (2021).
36. Lemos, P., Jeffrey, N., Cranmer, M., Ho, S. & Battaglia, P. Rediscovering orbital mechanics with machine learning. *Mach. Learn. Sci. Technol.* **4**, 045002 (2023).
37. Constantin, A. & Lukas, A. Formulae for line bundle cohomology on Calabi–Yau threefolds. *Fortsch. Phys.* **67**, 1900084 (2019).
38. Klaewer, D. & Schlechter, L. Machine learning line bundle cohomologies of hypersurfaces in toric varieties. *Phys. Lett. B* **789**, 438–443 (2019).
39. Altman, R., Carifio, J., Halverson, J. & Nelson, B. D. Estimating Calabi–Yau hypersurface and triangulation counts with equation learners. *J. High Energy Phys.* **03**, 186 (2019).
40. Grimm, T. W., Ruehle, F. & van de Heisteeg, D. Classifying Calabi–Yau threefolds using infinite distance limits. *Commun. Math. Phys.* **382**, 239–275 (2021).
41. Erbin, H. & Finotello, R. Machine learning for complete intersection Calabi–Yau manifolds: a methodological study. *Phys. Rev. D* **103**, 126014 (2021).
42. Berman, D. S., He, Y.-H. & Hirst, E. Machine learning Calabi–Yau hypersurfaces. *Phys. Rev. D* **105**, 066002 (2022).
43. Berglund, P., Campbell, B. & Jejjala, V. Machine learning Kreuzer–Skarke Calabi–Yau threefolds. Preprint at https://arxiv.org/abs/2112.09117 (2021).
44. Coates, T., Kasprzyk, A. M. & Veneziale, S. Machine learning the dimension of a Fano variety. *Nat. Commun.* **14**, 5526 (2023).
45. He, Y.-H. & Kim, M. Learning algebraic structures: preliminary investigations. Preprint at https://arxiv.org/abs/1905.02263 (2019).
46. Bao, J. et al. Quiver mutations, Seiberg duality and machine learning. *Phys. Rev. D* **102**, 086013 (2020).
47. Davies, A. et al. Advancing mathematics by guiding human intuition with AI. *Nature* **600**, 70–74 (2021).
48. Zaremba, W., Kurach, K. & Fergus, R. Learning to discover efficient mathematical identities. In *Proc. 27th International Conference on Neural Information Processing Systems* Vol. 1 (eds Ghahramani, Z. et al.) 1278–1286 (NeurIPS, 2014).
49. England, M. in *Mathematical Software – ICMS 2018* Vol. 10931 (eds Davenport, J. et al.) 165–174 (Springer, 2018).
50. Lample, G. & Charton, F. Deep learning for symbolic mathematics. In *Proc. Eighth International Conference on Learning Representations* (ICLR, 2023).
51. Weinan, E. Machine learning and computational mathematics. *Commun. Comput. Phys.* **28**, 1639–1670 (2020).
52. Peifer, D., Stillman, M. & Halpern-Leistner, D. Learning selection strategies in Buchberger's algorithm. In *Proc. 37th International Conference on Machine Learning* 7575–7585 (PMLR, 2020).
53. Kamienny, P.-A., d'Ascoli, S., Lample, G. & Charton, F. End-to-end symbolic regression with transformers. *Adv. Neural Inf. Process. Syst.* **35**, 10269–10281 (2022).
54. Dabelow, L. & Ueda, M. Symbolic equation solving via reinforcement learning. Preprint at https://arxiv.org/abs/2401.13447 (2024).
55. Ashmore, A., He, Y.-H. & Ovrut, B. A. Machine learning Calabi–Yau metrics. *Fortsch. Phys.* **68**, 2000068 (2020).
56. Jejjala, V., Mayorga Pena, D. K. & Mishra, C. Neural network approximations for Calabi–Yau metrics. *J. High Energy Phys.* **08**, 105 (2022).
57. Larfors, M., Lukas, A., Ruehle, F. & Schneider, R. Numerical metrics for complete intersection and Kreuzer–Skarke Calabi–Yau manifolds. *Mach. Learn. Sci. Technol.* **3**, 035014 (2022).
58. Alessandretti, L., Baronchelli, A. & He, Y.-H. Machine learning meets number theory: the data science of Birch–Swinnerton-Dyer. Preprint at https://arxiv.org/abs/1911.02008 (2019).
59. He, Y.-H., Lee, K.-H. & Oliver, T. Machine-learning the Sato–Tate conjecture. Preprint at https://arxiv.org/abs/2010.01213 (2020).
60. Abel, S. A. & Nutricati, L. A. Ising machines for diophantine problems in physics. *Fortsch. Phys.* **70**, 2200114 (2022).
61. Charton, F. Learning the greatest common divisor: explaining transformer predictions. Preprint at https://arxiv.org/abs/2308.15594# (2024).
62. He, Y.-H. & Yau, S.-T. Graph Laplacians, Riemannian manifolds and their machine-learning. Preprint at https://arxiv.org/abs/2006.16619 (2020).
63. Bao, J. et al. Hilbert series, machine learning, and applications to physics. *Phys. Lett. B* **827**, 136966 (2022).
64. Gukov, S., Halverson, J., Ruehle, F. & Sułkowski, P. Learning to unknot. *Mach. Learn. Sci. Technol.* **2**, 025035 (2021).
65. Craven, J., Jejjala, V. & Kar, A. Disentangling a deep learned volume formula. *J. High Energy Phys.* **06**, 040 (2021).
66. Cohen, T., Freytsis, M. & Ostdiek, B. (Machine) learning to do more with less. *J. High Energy Phys.* **2018**, 1–28 (2018).
67. Hudson, M. No coding required: companies make it easier than ever for scientists to use artificial intelligence. *Science* **365**, 416–417 (2019).
68. Xu, Y. et al. Artificial intelligence: a powerful paradigm for scientific research. *Innov. J.* **2**, 100179 (2021).
69. Georgescu, I. How machines could teach physicists new scientific concepts. *Nat. Rev. Phys.* **4**, 736–738 (2022).
70. Thiyagalingam, J., Shankar, M., Fox, G. & Hey, T. Scientific machine learning benchmarks. *Nat. Rev. Phys.* **4**, 413–420 (2022).
71. Gukov, S., Halverson, J. & Ruehle, F. Rigor with machine learning from field theory to the Poincaré conjecture. *Nat. Rev. Phys.* **6**, 310–319 (2024).
72. Fink, T. Why mathematics is set to be revolutionized by AI. *Nature* **629**, 505 (2024).
73. Bengio, Y. & Malkin, N. Machine learning and information theory concepts towards an AI mathematician. Preprint at https://arxiv.org/abs/2403.04571 (2024).
74. Satariano, A. & Specia, M. Global leaders warn AI poses risk of 'catastrophic' harm. *The New York Times* (1 November 2023).
75. Russell, S. in *Perspectives on Digital Humanism* (eds Werthner, H. et al.) 19–24 (Springer, 2022).
76. Lu, P., Qiu, L., Yu, W., Welleck, S. & Chang, K.-W. A survey of deep learning for mathematical reasoning. In *Proc. 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (eds Rogers, A. et al.) 14605–14631 (Association for Computational Linguistics, 2023).
77. Zhang, C. E., Collins, K. M., Weller, A. & Tenenbaum, J. B. AI for mathematics: a cognitive science perspective. Preprint at https://arxiv.org/abs/2310.13021 (2023).
78. Williamson, G. Is deep learning a useful tool for the pure mathematician? *Bull. Amer. Math. Soc.* **61**, 271–286 (2024).
79. He, Y.-H. (ed.) *Machine Learning in Pure Mathematics and Theoretical Physics* (World Scientific, 2023).
80. Wilkins, A. How AI mathematicians might finally deliver human-level reasoning. *New Scientist* (10 April 2024).
81. Greiner-Petter, A. et al. Why machines cannot learn mathematics, yet. Preprint at https://arxiv.org/abs/1905.08359 (2019).
82. Dolotin, V. V., Morozov, A. Y. & Popolitov, A. V. Machine learning of the well-known things. *Theor. Math. Phys.* **214**, 446–455 (2023).

83. Fajardo-Fontiveros, O. et al. Fundamental limits to learning closed-form mathematical models from data. *Nat. Commun.* https://doi.org/10.1038/s41467-023-36657-z (2023).
84. Kolpakov, A. & Rocke, A. On the impossibility of discovering a formula for primes using AI. Preprint at https://arxiv.org/abs/2308.10817 (2023).
85. He, Y.-H. Machine-learning mathematical structures. Preprint at https://arxiv.org/abs/2101.06317 (2021).
86. Lindström, S. et al. (eds) *Logicism, Intuitionism, and Formalism: What Has Become of Them?* Vol. 341 (Springer, 2008).
87. Zach, R. in *Philosophy of Logic* 411–447 (Elsevier, 2007).
88. Russell, B. & Whitehead, A. N. *Principia Mathematica* 2nd edn, Vol. 1 (Cambridge Univ. Press, 1963).
89. Ioannidis, J. P., Fanelli, D., Dunne, D. D. & Goodman, S. N. Meta-research: evaluation and improvement of research methods and practices. *PLoS Biol.* **13**, e1002264 (2015).
90. McCarthy, J. in *Daedalus* Vol. 117, 297–311 (MIT Press, 1988).
91. Brouwer, L. E. J. in *From Frege to Gödel: A Source Book In Mathematical Logic 1879–1931* (ed. van Heijenoort, J.) 490–492 (Harvard Univ. Press, 1967).
92. Gödel, K. Über formal unentscheidbare sätze der principia mathematica und verwandter systeme I [German]. *Monatshefte für mathematik und physik* **38**, 173–198 (1931).
93. Church, A. An unsolvable problem of elementary number theory. *Am. J. Math.* **58**, 345–363 (1936).
94. Turing, A. M. On computable numbers, with an application to the Entscheidungsproblem. A correction. *Proc. Lond. Math. Soc.* **s2-43**, 544–546 (1938).
95. Newell, A. & Simon, H. The logic theory machine — a complex information processing system. *IRE Trans. Inf. Theory* **2**, 61–79 (1956).
96. Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **65**, 386 (1958).
97. Roberts, S. AI is coming for mathematics, too. *The New York Times* (2 July 2023).
98. Davenport, J. & Buzzard, K. The future of mathematics. In *Proc. International Congress of Mathematicians* (eds Sirakov, B. et al.) 1085–1110 (World Scientific, 2018).
99. De Bruijn, N. G. in *Studies in Logic and the Foundations of Mathematics* Vol. 133, 73–100 (Elsevier, 1994).
100. Nipkow, T. et al. (eds) *Isabelle/HOL: A Proof Assistant for Higher-order Logic* Vol. 2283 (Springer, 2002).
101. Bertot, Y. & Castéran, P. *Interactive Theorem Proving and Program Development: Coq'Art: The Calculus of Inductive Constructions* (Springer, 2004).
102. Wadler, P. Programming language foundations in Agda. In *Formal Methods: Foundations and Applications (SBMF 2018)* (eds Massoni, T. & Mousavi, M.) 56–73 (Springer, 2018).
103. de Moura, L., Kong, S., Avigad, J., van Doorn, F. & von Raumer, J. in *Automated Deduction-CADE-25: 25th International Conference on Automated Deduction, Berlin, Germany, August 1–7, 2015, Proceedings* (eds Felty, A. P. & Middeldorp, A.) 378–388 (Springer, 2015).
104. Gowers, W. T., Green, B., Manners, F. & Tao, T. On a conjecture of Marton. Preprint at https://arxiv.org/abs/2311.05762 (2023).
105. Ganesalingam, M. *The Language of Mathematics* (Springer, 2013).
106. Wittgenstein, L. *Tractatus logico-philosophicus* (Kegan Paul, Trench, Trübner & Co., 1922).
107. Turing, A. M. Computing machinery and intelligence. *Mind* **LIX**, 433–460 (1950).
108. Biever, C. ChatGPT broke the turing test — the race is on for new ways to assess AI. *Nature* **619**, 686–689 (2023).
109. He, Y.-H., Jejjala, V. & Nelson, B. D. hep-th. Preprint at https://arxiv.org/abs/1807.00735 (2018).
110. Mikolov, T., Chen, K., Corrado, G. & Dean, J. Efficient estimation of word representations in vector space. In *Proc. 1st International Conference on Learning Representations* (ICLR, 2013).
111. Tshitoyan, V. et al. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature* **571**, 95–98 (2019).
112. Kim, J. T., Landajuela, M. & Petersen, B. K. Distilling Wikipedia mathematical knowledge into neural network models. In *Proc. 1st Mathematical Reasoning in General Artificial Intelligence Workshop* (ICLR, 2021).
113. Hutson, M. AI learns to write computer code in 'stunning' advance. *Science* https://doi.org/10.1126/science.adg2088 (2022).
114. Wang, Z., Xia, R. & Liu, P. Generative AI for math: Part i — MathPile: a billion-token-scale pretraining corpus for math. Preprint at https://arxiv.org/abs/2312.17120 (2023).
115. Romera-Paredes, B. et al. Mathematical discoveries from program search with large language models. *Nature* **625**, 468–475 (2024).
116. Trinh, T. H., Wu, Y., Le, Q. V., He, H. & Luong, T. Solving olympiad geometry without human demonstrations. *Nature* **625**, 476–482 (2024).
117. Ahn, J. et al. Large language models for mathematical reasoning: progresses and challenges. Preprint at https://arxiv.org/abs/2402.00157 (2024).
118. Frieder, S., Berner, J., Petersen, P. & Lukasiewicz, T. Large language models for mathematicians. *Internationale Mathematische Nachrichten* **254**, 1–20 (2024).
119. Didolkar, A. et al. Metacognitive capabilities of LLMs: an exploration in mathematical problem solving. Preprint at https://arxiv.org/abs/2405.12205 (2024).
120. Polu, S., Han, J. M. & Sutskever, I. Solving (some) formal math olympiad problems. *OpenAI* https://openai.com/index/formal-math/ (2022).
121. Rodriguez, A., Azhar, F. & Spisak, J. Teaching AI advanced mathematical reasoning. *Meta* https://ai.meta.com/blog/ai-math-theorem-proving/ (2022).
122. Otte, M. & Panza, M. (eds) *Analysis and Synthesis in Mathematics: History and Philosophy* Vol. 196 (Kluwer Academic Publishers, 1997).
123. Lynch, P. Will mathematicians be replaced by computers? *The Irish Times* https://www.irishtimes.com/news/science/will-mathematicians-be-replaced-by-computers-1.4349668 (17 September 2020).
124. Penrose, R. The role of aesthetics in pure and applied mathematical research. *Bull. Inst. Math. Appl.* **10**, 266–271 (1974).
125. Arnol'd, V. I. On teaching mathematics. *Uspekhi Mat. Nauk* **53**, 229–234 (1998).
126. Carlson, J. A., Jaffe, A. & Wiles, A. (eds) *The Millennium Prize Problems* (American Mathematical Society, 2006).
127. Hardy, G. H. *A Mathematician's Apology* (Cambridge Univ. Press, 1992).
128. Wu, C. W. Can machine learning identify interesting mathematics? An exploration using empirically observed laws. Preprint at https://arxiv.org/abs/1805.07431 (2018).
129. Friederich, P., Krenn, M., Tamblyn, I. & Aspuru-Guzik, A. Scientific intuition inspired by machine learning-generated hypotheses. *Mach. Learn. Sci. Technol.* **2**, 025027 (2021).
130. Raayoni, G. et al. Generating conjectures on fundamental constants with the Ramanujan machine. *Nature* **590**, 67–73 (2021).
131. Mishra, C., Moulik, S. R. & Sarkar, R. Mathematical conjecture generation using machine intelligence. Preprint at https://arxiv.org/abs/2306.07277 (2023).
132. Bauer, A., Petković, M. & Todorovski, L. MLFMF: data sets for machine learning for mathematical formalization. In *Proc. 37th International Conference on Neural Information Processing Systems* 50730–50741 (Association for Computing Machinery, 2024).
133. Davila, R. Advancements in research mathematics through AI: a framework for conjecturing. Preprint at https://arxiv.org/abs/2306.12917 (2023).
134. Birch, B. Reminiscences from 1958–62. *Isaac Newton Institute for Mathematical Sciences* https://www.sms.cam.ac.uk/media/3530787 (2021).
135. He, Y.-H. & Burtsev, M. Can AI make genuine theoretical discoveries? *Nature* **625**, 241–241 (2024).
136. Schenck, H. *Computational Algebraic Geometry* Vol. 58 (Cambridge Univ. Press, 2003).
137. Fawzi, A. et al. Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature* **610**, 47–53 (2022).
138. Kauers, M. & Moosbauer, J. The FBHHRBNRSSSHK-algorithm for multiplication in $\mathbb{Z}_2^{5\times 5}$ is still not the end of the story. Preprint at https://arxiv.org/abs/2210.04045 (2022).
139. He, Y.-H., Lee, K.-H., Oliver, T. & Pozdnyakov, A. Murmurations of elliptic curves. Preprint at https://arxiv.org/abs/2204.10140 (2022).
140. Gorenstein, D. *Finite Simple Groups: An Introduction to Their Classification* (Springer, 2013).
141. Touvron, H. et al. LLaMA: open and efficient foundation language models. Preprint at https://arxiv.org/abs/2302.13971 (2023).
142. Wolfram, S. ChatGPT gets its 'Wolfram superpowers'! *Stephen Wolfram Writings* https://writings.stephenwolfram.com/2023/03/chatgpt-gets-its-wolfram-superpowers/ (2023).
143. Ellenberg, J. Will AI replace mathematicians? *Big Think* https://bigthink.com/the-future/artificial-intelligence-replace-mathematicians/ (2021).
144. Qin, H. & Ye, Y. Algorithms of the Möbius function by random forests and neural networks. *J. Big Data* **11**, 31 (2024).
145. Mossinghoff, M. J. & Trudgian, T. S. in *Exploring the Riemann Zeta Function: 190 Years from Riemann's Birth* (eds Montgomery, H. et al.) 201–221 (Springer, 2017).

Acknowledgements

The author is most grateful to A. Bhattacharya, A. Kosyak and M. Duncan for many valuable comments on the draft. The author thanks many collaborators over the past few years on AI-assisted mathematics, for the great fun and friendship: D. Aggarwal, L. Alessandretti, G. Arias-Tamargo, A. Ashmore, J. Bao, A. Baronchelli, P. Berglund, D. Berman, K. Bull, L. Calmon, H. Chen, S. Chen, A. Constantin, P.-P. Dechant, R. Deen, S. Garoufalidis, E. Heyes, E. Hirst, J. Hofscheier, J. Ipiña, V. Jejjala, A. Kasprzyk, M. Kim, S. Lal, K.-H. Lee, S.-J. Lee, J. Li, A. Lukas, S. Majumder, C. Mishra, G. Musiker, B. Nelson, A. Nestor, T. Oliver, B. Ovrut, T. Peterken, S. Pietromonaco, A. Pozdnyakov, D. Riabchenko, D. Rodriguez-Gomez, H. Sá Earp, M. Sharnoff, T. Silva, E. Sultanow, Y. Xiao, S.-T. Yau and Z. Zaz. The research is funded in part by STFC grant ST/J00037X/2 and the Leverhulme Trust for a project grant.

Competing interests

The author declares no competing interests.

Additional information

Peer review information *Nature Reviews Physics* thanks Rak-Kyeong Seong, Adam Zsolt Wagner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.