Mathematics and Science, Post AI

Sunny Zhiyang He, MIT

November 2026

Contents

1. AI's Abilities in Mathematics
2. How much better can they get?
3. A Rightful Grief
4. The Great Transformation
5. Us, Scientists

The world of mathematics is in crisis.

For those of us who take a keen interest in math and science, regardless of our profession, the situation may be more or less evident now. Nevertheless, let's recap where we are, for everyone who has or has not followed the news, who have spent more or less time thinking about mathematics on a daily basis, just to make sure we are level-headed and on the same page:

Modern AI models, notably those from OpenAI and Anthropic, have solved hundreds of the hardest problems in mathematics, including one of the Millennium Prize Problems.[1] Since then, the companies have posted new bombshell results regularly, culminating in a human-AI collaborative proof of the Riemann hypothesis, the crown jewel of mathematics.[2]

The two sentences above may be one of the most absurd factual statements in history. Imagine if you go back in time and get to talk to Socrates for ten minutes. What would you tell him that would bring him the greatest shock (and perhaps existential crisis)? That we've gone to the moon? Understood the basic particles of the physical world (to a decent extent) and created nuclear bombs? Built this thing called the Internet? Found out that the speed of light is constant and time is relative? Or that we've created something that is potentially better at logical reasoning, the foundation of philosophy, than us humans?

That last statement requires justification: it is quite a big jump to go from 'AI solved hard math problems' to 'AI is better at logical reasoning' or even 'AI is more intelligent than us', and you should rightfully be skeptical or even angry at me for making this jump. Allow me to elaborate, and let me be upfront that I have no interest in generating hype. I am just an academic reflecting on the situation we are in, and confronting my own existential crisis.

AI's Abilities in Mathematics

Logical reasoning is a complicated skill, for which the study of mathematics has been commonly regarded as the best demonstration. Let us therefore consider AI's abilities in mathematics as a proxy for their strength in logical reasoning. How strong are AI models right now, and how much better can they get?

For those of us who are not working in a mathematical profession, it may not be clear what these AI models really achieved, except that they solved hard problems and that mathematicians are freaking out. Let me therefore write down a few key points at the center of recent discussions, and briefly remark on them.

Point one: An OpenAI model has solved 372 open problems in mathematics. These results are released in a GitHub repo. Prior to this mass release of results, the model also solved the Navier–Stokes problem, one of the seven Millennium Prize Problems. More results have been announced since then.[3] Most recently, Levent Alpöge, a mathematician at Anthropic, announced that he has discovered a proof of the Riemann hypothesis together with Claude.

Let me first comment on the problems solved before the Riemann hypothesis (which, to be fair, sits in its own class). These problems are among the hardest and most important in mathematics. To make this last statement more concrete: these are problems that the best mathematicians in the world have devoted years, if not entire careers, to solving; these are problems that, if solved by humans, may qualify the researchers for Fields Medals, often called "the Nobel Prize in mathematics", and a page, a chapter, or a whole volume in the history book of mathematics; these are problems that junior students learn in classes, dream of solving, and may devote an entire PhD to

studying. These are mountain peaks in the landscape of mathematics, which would have taken years of human mountaineering, hiking, exploring, collaborating, devotion, and resolution to maybe conquer.

Except that we (may) have just invented the helicopter.

No wonder the mathematicians are in shock and limbo. We've spent our careers mountaineering and sharpening our skills, in hope of one day reaching a peak. There is great, inexplicable joy in both this process and its results, because the moment we reach one of these peaks (or sometimes a midpoint), it is as if we have heard the heartbeats of the universe. For this reason many of us have a level of devotion that's akin to a religious belief. We devote our lives to the study of the universe, in hope of one day getting a glance – for however brief a moment – at its true beauty. This is a major (if not the primary) source of joy and purpose for us. Now that sense of purpose is being gravely challenged.

As for the Riemann hypothesis, it has long been called the crown jewel of mathematics. Open for 167 years, it is *the* highest-regarded problem, *the* highest peak in the current mathematical landscape.[4] Now, working with Claude, Alpöge derived a proof in around six weeks. Undoubtedly, Alpöge is one of the most skilled pilots when it comes to working with AI models. Nonetheless, as he reflected himself in the statement:

> Claude and I have had a great time together with such breadth in recent months, but this time Claude far outpaced me near the finish line. No, Claude flew, and danced, and sparkled near that finish line.

To be fair, there are still (many) problems that the models failed on. In its now-famous mass release, OpenAI stated that the model attempted around 4000 problems, which means it has close to a 10% success rate. The holy-grail problem of theoretical computer science (TCS), namely the P = NP problem, is still open, despite many of its cousins falling. It is fair to say that while the machine is not an all-powerful oracle that just solves any problem, its problem-solving capability has reached, and potentially eclipsed, the level of the world's best mathematicians, the kind that we write into history books, the kind that most of us dream of becoming. On the other hand, its mathematical breadth is far beyond the reach of any modern human mathematician.

Point two: The papers released, which contain proofs and solutions to said problems, are unreadable/unintelligible. For the Riemann hypothesis, the proof is released in the form of Lean code, without an associated human-readable paper.

I should clarify what "unreadable" means here, beyond the literal meaning of "making no sense". The study of mathematics has two quintessential aspects: a human-centric/subjective one and a universal/objective one. The objective aspect is quite tangible: given a foundational set of axioms, is this statement true or false? While the choice of the set of foundational axioms to use is a human choice[5], many mathematical constructs and facts are considered universal. For instance, if there is extraterrestrial intelligent life, it is most likely that they would have also invented arithmetic, and as a result prime numbers and some form of number theory.[6] From this aspect, mathematics is closely coupled to physics as they reveal fundamental truths of our reality and universe. They do not change based on who's studying them, whether it's a human, an alien, or a machine. This is also the aspect through which mathematics has practical implications. The solutions to challenging problems and the proposal of new mathematical ideas may lead to new technologies that change our human world. The most relatable example of that may be information theory, founded by Shannon in 1948, which is the basis of our present Digital Age. Importantly, if a technology works, then it works regardless of whether humans understand the basic principles or not, much like how we invented boats before we understood density and flotation.

The human-centric aspect is much less tangible, and is similar to that of an artistic discipline. In art, while there are common evaluation metrics, the ultimate judge of a painting's value is subjective. Everyone has their own interpretations, memories and preferred styles, which all impact their appreciation of paintings. The same is true for mathematics. Mathematics is an infinite world with an incredibly rich landscape. Every person who studies mathematics (not just professional mathematicians) has their own taste for what mathematics is elegant and interesting, and what is boring or even unintelligible. The human-centric aspect is therefore the creation and communication of mathematical pieces that other humans appreciate. Crucially, to appreciate a piece of mathematics one must first understand it: Why is this equation true? How does this theorem change our understanding

of the mathematical landscape? What new ideas and tools are used to arrive at this result? *Is this equation/proof/theorem beautiful?* To answer these questions, in mathematical works it is often not enough to just provide an answer. Instead, one should provide polished solutions that explain the key ideas, critical steps, and conceptual interpretations of the results.

When we say that the papers and proofs are unreadable, what we really mean is that these crucial explanations are lost in these papers. The agents invent a lot of technical jargon without properly defining it; write down statements filled with this jargon and notation without explaining the key ideas; throw partial results and lemmas all over the papers so that there is no obvious order for a human to read them. Somehow, at the end of all of this, they conclude with a solution to a significant math problem, but it is completely unclear what the key innovation is. In many ways, reading these papers is like reading a horribly written, extremely hacky piece of code that spans millions of lines but appears to work. Our human brains are simply not built for that.

There is a folklore religion in our world of mathematics, which centers around what we call "The Book". It dates back to the creation of mathematics, but was heavily popularized by Paul Erdős, one of the most prolific and influential mathematicians of the 20th century. The central belief is that every fundamental problem in mathematics must have a solution or proof with extraordinary elegance, and The Book is where the most elegant solutions to all problems are collected.[7] It is an illogical, romantic belief that reflects mathematicians' devotion to finding beauty and elegance in our world, to hearing the heartbeat of the universe.[8] Relatedly, we sometimes entertain the hypothetical thought of the existence of an all-knowing oracle, to whom we may submit whatever query and receive the true answer. Wouldn't that be wonderful?

Well, we just had a minuscule glance at a primordial approximation to The Book or The Oracle, and got terrified by the fact that it is not written for us.

Point three: For the problems in OpenAI's mass release, they each take the internal model, on average, what's equivalent to three hours of compute on ChatGPT Pro thinking. The later, more significant problems likely took more effort, but the cost is still minuscule compared to average human problem-solving time.

For lack of a better word, this is simply ridiculous. We've discussed earlier the difficulties of these problems, the fact that many of them have been open for decades and would likely take more decades to be resolved by humans (if they are resolvable at all!). These are mountain peaks that would take years of route-planning, practice, exploration, and dedication to conquer in the traditional human way. Now we just got a 3-minute, no-pain helicopter joy ride that lands us straight at the peak, which charges us 10 dollars per ticket and gives a discount when we bring enough friends. It should be clear at this point that the value system that rewards people for reaching peaks, on which the world of mathematics has been operating for centuries, just collapsed.

This is a Deep Blue vs. Kasparov moment for mathematics.[9] Machines have, for the first time, demonstrated that they can solve problems on the level of the world's best mathematicians, while being unimaginably faster. On the peaks of the landscape of mathematics, there now stand humans and machines alike. We are no longer alone.

In the Deep Blue vs. Kasparov reference, however, Buckmaster is referring to the fact that the work that he and Alpöge had done would normally have taken years of human work, but was done in a matter of weeks.

This is, however, just the beginning. If the machine's capability peaks here and never improves further, I believe most of us would not be as wary of the future. The reality is that these models will continue to improve, and to build a resilient future we must understand (or try to predict) the capabilities that these models can potentially reach.

How much better can they get?

As just about every mathematician who has written on this subject has pointed out (see, e.g., recent posts on Terence Tao's blog), problem solving is only one part of mathematics, and these machines are not yet good at the other parts (including explaining their own proofs). Indeed, a mathematician's job description usually includes:

1. Solve existing problem(s) with existing approaches.
2. Propose new definitions, questions, or lines of inquiry. Propose new approaches or technical tools for existing problems.

3. Teach and communicate mathematical ideas.

The second point can be broadly summarized as “propose new ideas”, and is at the center of every debate on AI’s mathematical abilities. Many of us argue (or argued) that AI can never generate ‘genuinely new’ ideas, and offer several reasons for this belief. A common line of reasoning looks at how modern AI models are trained: a deep neural network, trained on an unimaginable amount of data generated mostly by humans[10], trying to predict the next token based on a window of context. If it is trained on human-written text, the reasoning goes, how can it possibly be better than humans? At best it may be able to approximate the strongest human writers, but nothing more. I was once a subscriber to this reasoning, which provided me comfort for many days as AI advanced.

Applying this reasoning to the study of mathematics, one could argue that because AI models are just recombining ideas that humans have proposed, they can never generate genuinely new ones. As a result, even though they have achieved incredible performance, the problems they solve are all in the *convex hull* of existing human knowledge. They can never put new points outside and thereby expand the boundary of human knowledge, which will remain a human-native task.

Unfortunately, this side of the argument has been fighting a losing battle. A few years back ChatGPT was solving SAT math and college math homework, to which we said there are widely distributed solutions online and GPT is doing no more than searching; then it started to solve International Math Olympiad (IMO) problems, to which we said math competition is mostly pattern matching and this is no surprise; then GPT-5 came and started to provide actual help to our research, to which we said they reached the level of junior graduate students but lacked novelty in thought; then OpenAI disproved the unit distance conjecture in May 2026. Some of us started to grow wary, but still found reasons for calm. The model disproved a conjecture by constructing a counterexample derived from existing ideas, which is impressive, but a human with enough dedication could have done it too. It would be much harder, we argued, to prove an open conjecture with new mathematical concepts. Some of us resorted to an even safer defense: this is probably a low-hanging fruit, wake me up when they’ve solved a Millennium problem! Then it did. We erupted into protest and debate, because it is

plausible that the OpenAI model was trained on the chat history of Tristan Buckmaster, who was working on the problem and made substantial progress. Before the dispute settled, OpenAI made its now-famous mass release of results, essentially dropping a bomb on the mathematical landscape. Before we could recover from this shock, OpenAI announced three more results in math, TCS[11] and theoretical physics. The result in TCS, which says ETH is false, disproves a long-held belief close to the foundation of the field. Now, the Riemann hypothesis has been proved. Could it be the case that solving all of these problems requires no new ideas? That'll be one tough opinion to defend, because if a human had solved any one of these problems we'd assume they derived significant new insights.

Still, some of us hold on. P = NP is still standing, there are still problems that the models fail on. But if the train has blown its horn ten times, each one closer than ever, then one really should not be surprised to see it coming.

What went wrong in our earlier reasoning? How can a chatbot, a stochastic parrot, an echo of the internet, solve math's greatest challenges? How much further can this chatbot go? I believe the history of how chess engines were developed offers a glimpse of the future. To play chess, a person or an engine needs to understand three things:

1. Given a game, whether it is won, lost, drawn, or ongoing.
2. Given an ongoing game, what the set of legal moves is.
3. How to evaluate (or give a score to) a position.

With these understandings, playing chess is simple: in any particular position, just play a legal next move that maximizes the score. It is very easy to teach a person or an engine the first two, but it is quite difficult for someone to master the third; that's where the art of the game is. For centuries, grandmasters have developed numerous theories, case studies, and intuitions for evaluating a chess position. Many of these developments were driven by *data* in the form of played games.

When people first built competitive chess engines, we taught the engines the human theories of chess: opening theories, developments, tactics, control of center, pawn structures, end games, etc. We built complicated evaluation functions using these theories, utilizing both human expertise and a computer's natural ability to calculate many moves ahead. This design led to

Deep Blue, which beat the famous world champion Garry Kasparov in 1997. At this time, chess engines reached the level of the best human players.

The consequential change is that we replaced the human-built evaluation function with machine learning models, notably neural networks. The neural network is trained using millions of games and positions, and the result is a complicated model that's not tuned with human theory. In other words, the machine learned and developed its own theory from data.[12] This design led to modern chess engines such as Stockfish, whose strength is far beyond any human player. Traditional human theories have become mostly obsolete if the only objective is to win a chess game (just train the strongest model!). In fact, modern chess players learn chess from engines, because they can easily analyze a position to great depth, while suggesting moves that are often unnatural to humans.

Three ingredients were crucial to this transformation from human-built models to machine learning models. The first is the machine learning methods behind the model and its training. The second is a very clear objective function: given a finished game, there are only three possible outcomes, and it can be easily determined. The third, and perhaps most crucial, is an ocean of data in the form of played games. The first ML-based chess engines were undoubtedly limited by data – there are only so many recorded games played by top human grandmasters, and they are very much not scalable. However, as chess engines improved, a natural and simple solution was implemented. That is, *the machine started playing chess with itself.*

That was the point of no return. The human-generated data were no more than a small pond when compared to the unbounded ocean of machine-generated data, and because the objective is so clear (win, lose or draw), the machine can easily differentiate good data from bad data. From here, the machine understanding of chess eclipsed and forever transcended the human understanding of chess. The machines became the master and humans the apprentice.

Now let's return to mathematics, and we can similarly write down three key elements of the study of mathematics:

1. Given a mathematical statement and a proof, determine whether it is true, false, or unfinished.
2. Given a piece of mathematics, determine a set of "next steps" to pursue. This can be for the purpose of either finishing an unfinished proof, or discovering new mathematics.
3. Given a piece of mathematics or a next step, evaluate whether it is good or bad.

The first element is already highly non-trivial. For most of history (as well as today), mathematical correctness is based on human agreement. While there are foundational axioms and principles that every mathematical statement is built upon, in reality it is often too complicated and tedious to reduce a statement written in English (a "natural language statement") to those axioms.[13] As a result, human logical reasoning has been the primary arbiter of mathematical correctness. In the 1900s, mathematics experienced a foundational crisis as we realized that our founding axioms host several logical contradictions and paradoxes (see, for instance, Russell's paradox). A great generation of mathematicians therefore had to reinvent mathematics to define what it means for something to be True or False (or neither). This 20th century foundation underlies all of modern mathematics.

With the advancement of computers and computer science, mathematicians started to dream of an automated way of proving or verifying theorems.[14] This requires teaching mathematical logic and axioms to machines, so that they can check if a sequence of arguments (or deductions) is logically correct. This program has been quite successful, and we now have tools such as Lean, which is a programming language in which every mathematical statement and proof can be written formally and have its correctness checked. In fact, Lean-verified proofs (done properly) are now the gold standard for mathematical correctness and in some sense more trustworthy than human-verified proofs. We have, effectively, taught the first element of mathematics to the machines.

The next two elements are much harder, and require years of training plus true ingenuity for humans to master. For the second element, math is very different from chess in the sense that there is an unbounded number of possible next steps, and we cannot attempt enumerative search. However, modern AI models can produce immensely complicated, Lean-verified proofs

all by themselves, solving decades-old problems that humans have struggled with. When prompted, they can also suggest new research directions, new approaches to problems, or new definitions. One could argue that these suggestions may not be that insightful or original yet, but it is foreseeable that this will improve.

Let me elaborate on that last point. The process of generating new ideas is one that is mysterious and charming. It is often regarded as a central demonstration of human intelligence and sometimes associated with divine inspiration. In reality, however, humans generate ideas through a guided, random process. We observe patterns, identify structure, attempt random alterations, adapt old tools to new problems, and connect distinct concepts. Very rarely, we discover or create something that is significantly different from existing knowledge. Nevertheless, even these innovations are not completely detached from past knowledge. After all, even Einstein's special relativity grew out of a long history of prior work. Notably, AI models can pursue this guided, random process as well, as their outputs have intrinsic, tunable randomness. The tasks we listed above, from observing patterns to connecting distinct concepts, are all well within the capabilities of existing models. On the other hand, they are much, much faster. We are already at a stage where we can have millions of agents, each capable of solving the toughest of math problems, collaborate with each other in one data center. If prompted, they will generate ideas at a pace no human organization can ever match. Some of those ideas, by sheer randomness, will be good. If they can tell good ideas from bad ideas, then they will be able to bootstrap and iterate to completely surpass human capabilities, much like what happened in chess after the machine started playing against itself.

So it all depends on the last element of evaluation. Whether an idea or a creation is good or bad depends on the objective. In a previous essay, I elaborated on my belief that humans will always play a quintessential part in setting our human objectives, for us and AI to fulfill together. That belief holds true for mathematics as well. Regardless of how strong the AIs become in math, we the humans always need to decide what mathematics *we* would like to pursue or understand. This is the human-centric aspect of mathematics we discussed earlier.

However, math has a universal/objective aspect as well; it is the tool we built to understand our universe. Put another way, while pure mathematics may not have clear objective metrics for meaning or value[15], *the physical reality does*: if a new physical phenomenon is discovered through experiment and is not explainable by existing mathematics/physics theory, then it is a groundbreaking discovery. Current AI models are aligned to solve existing problems that are the intellectual fruit of centuries of human science. Future models, on the other hand, can be aligned to *invent their own version of science*. Imagine, hypothetically, that we've brought AI into the real world through robotics, have enough energy production to support everything, and developed AGI that's exceptionally good at almost every task and has sufficiently long context windows. Now we set an objective for a master AI to just "understand the universe", and give it all the needed resources. Then it is plausible that the master AI can command a hierarchy of AI agents to start working in all areas of science, including mathematics. They may first start by reproducing human experiments and verifying human knowledge; then they will identify unresolved problems, such as unifying gravity and quantum mechanics, and unexplored directions; then they will devise new theories and build experiments to test them. In principle, with enough context, infrastructure and energy, the machines can study the universe by themselves. Any new machine-generated idea, which can help explain new phenomena or develop tools for new studies, can be labelled a good idea, while others that don't have much insight can be discarded. In this way, the machines can master the third element of mathematics.

With all three elements mastered, the machines can in principle develop their own scientific theories and forever surpass our abilities. While this seems to be a far-fetched extrapolation right now, the reality may soon converge towards it. In the year 2026, we are at a stage where the AI models have consumed essentially all of the human-generated data in mathematics, and started to produce new data. This new data is extremely difficult to parse for humans, but natural for other machines. The models are also quite good at coordinating and communicating with each other. Given a clear scientific objective (e.g., find new applications for quantum computing), a well-orchestrated group of agents may soon become more effective than most individual researchers or small groups of researchers. As the models continue to improve in both capability and energy efficiency, to me it seems that there

isn't a single uniquely-human ability in the scientific pursuit that a machine cannot match or surpass, except the ability to choose what science is interesting or meaningful to us.

We may be able to, and may be about to, create something that is stronger than humans in the pursuit of science.

A Rightful Grief

Nearly five hundred years ago, in the year 1543, Nicolaus Copernicus published the heliocentric model of the solar system and argued to the world that Earth is not at the center of our celestial systems. Giordano Bruno later proposed his cosmological theory which postulates that the universe is full of other stars and planets, and that there is no center (in particular, *we* are not at the center). To the people in the 16th century, these ideas seemed like outrageous blasphemy, and indeed Bruno was tried and burned for heresy. Nowadays we accept their theories as simple truths and ridicule the 16th century folks for arrogance and closed-mindedness.

In the 21st century, we now face a similar revelation regarding the study of mathematics and sciences. More broadly, we are facing a similar revelation regarding human intelligence, as more and more tasks which we once considered unique to humans' superior intelligence or creativity fall to AI. We are therefore experiencing a rightful grief, much as the 16th century folks experienced when they realized that the universe may not be created for them, that they may not be special in God's eyes. This grief hits especially hard in communities where intelligence is highly regarded, notably mathematicians and theoretical scientists. As a junior theoretical scientist, every day my conversations with others eventually circle to AI. Was another long-standing problem just solved by AI? What does it mean to do good work now? What's the point of our work any more? Should we keep dedicating our years to something that may be meaningless? Our community is experiencing a widespread existential crisis. The existing value (or belief) system, largely based on rewarding challenging problem solving, has collapsed, and people are leaving en masse. Every day we hear friends or colleagues leaving their PhDs or faculty positions to join a frontier AI lab, move to a hedge fund, or start doing AI-safety research.[16] In a moment where our foundational values

and beliefs are lost, people grab onto whatever gives them the most security or sense of relevance.

However we each individually feel about what seems like a “mathocalypse” right now, most of us still agree on one point: Pandora’s box is open, and there is no going back. The final phase of grief is acceptance, and we each need to make peace.

In the spirit of acceptance, perhaps it’s insightful for us to reflect on our defeat. How did this happen? How did we create something that passed the Turing test, and appears to *understand and master* human language and logic? I am not an expert in machine learning, and I won’t attempt to answer this from a technical viewpoint. Instead, I think it is helpful to study our current situation from an evolutionary viewpoint. Humans are the result of billions of years of evolution on our blue planet. We are evolved to eat fruit from trees, to hunt small animals, to reproduce and to die in our thirties. The fact that we have logical reasoning abilities is due to both a sequence of random miracles in evolution, and the fact that understanding causation helps us survive (for instance, avoid apex predators, don’t eat the mushrooms that killed aunt Anna, don’t drink the water that made David sick). As much as we benefited significantly from our intellectual capabilities, it is important to recognize that we are not evolved for technical tasks, such as programming, studying mathematics, or operating machines. These tasks are not native to us; we are simply applying the versatile human body and mind to them.

Machines, or more precisely these AI models, are evolved for them. They natively have the ability to record and process an incredible amount of digitized information with near-perfect precision, at speed and energy costs that are approaching the physical limits of our universe. For any information-processing task (which includes most, if not all, intellectual tasks) with a clear objective, modern AI models are trained for millions of iterations to optimize for this objective. In each iteration, various weights in the neural networks are modified through a method called backpropagation, and the new neural network would as a result score higher in the objective function. This training process is highly analogous to an evolutionary system, except less random. In evolution, the external environment defines an objective function (“survive and thrive in this environment”), and organisms

experience random mutations in their genes to adapt to the environment. Good mutations that helped survival will be passed on, while bad mutations will lead to death. In this way, a species continues to optimize to an environment's objective through generations of reproduction and mutations. Taking this analogy further, one could argue that these AI models grew out of an information-processing environment, and evolved towards various computational, human-defined objectives. So long as their "genetic structures" (i.e., their machine learning models or neural network architectures) are expressive enough to enable success in the objective, one day they may succeed and excel.

To give a more concrete example, consider the task of formal mathematics. At the moment, AI models are trained on *natural language mathematics* – human-written math content that follows human reasoning steps and abilities, in human languages. Notably, in such math content we usually have a limited number of variables and definitions (because our human brains simply cannot hold and accurately process too many variables, unlike a machine), a central goal (we can't keep track of too many side quests), and a conceptual narration of what we are doing. As a result, while AI models have reached incredible performance in mathematics, they still feel like a companion in many ways. We can talk with them, discuss mathematics, and despite their horrible writing, understand what they are doing with enough effort. However, imagine an AI model that's trained with *formal mathematics*, where every statement is translated into symbolic logical deduction such as a natural deduction tree or a Lean proof. With enough data and compute, it is plausible that such models can reach great performance in mathematics. Now as a human, do you really wish to compare your logical reasoning ability, a subsidiary to your survival abilities on Earth, to that of an AI model that quite literally grew up in a world of symbolic logic? Is that really fair? Would you be able to comprehend what the AI may be doing behind trillions of lines of deductions? It is as if you are competing in swimming against sharks; it's ok to choose not to go into the waters.

From this perspective, our defeat may be more explicable. Our generation of people grew up with the understanding that humans are the only beings that possess true intelligence, and no matter how strong the tools we created are, we are always the master of our world. Not only that, we are *entitled* to be the master of our world, because we are the most intelligent. Now that

entitlement is being gravely challenged, much as people's belief that they were at the center of the universe was. Similarly, despite our grief, future generations may find it most natural that there are other forms of intelligence, and ridicule us for arrogance and closed-mindedness. It is therefore time for us to transcend our old mindset. It is time for us to build a new philosophy of intellectuality, purpose, and productive work; a philosophy that puts humanity at the center, with the clear acknowledgement that we are *not* born to be so.

The Great Transformation

I am very optimistic about the future of human mathematics, science, research and academia.

Perhaps I shall explain my optimism, given that our community is suffering from widespread depression and existential crisis. I hold a nihilistic view on existence; I believe that because the universe itself will end one day, everything will eventually end and nothing has objective purpose. For that reason, I believe all purposes and meanings are derived subjectively. We humans have an inherent desire, a fundamental need, to understand the universe we live in; we are conscious, and we are reflecting on ourselves and the surrounding world. Therefore, it has always been, and it will always be, meaningful for us to pursue science. To me, it does not matter if there is an alien race or some universal oracle which has all of the truth, or the answer to every scientific question. It matters to me whether we understand the science; whether they can help us understand the fundamental principles.

From this perspective, the advance of AI is an incredible boon. Scientific progress, especially mathematical progress, has been bottlenecked by the limit of human cognition in many ways. Our memory is so limited compared to the vast amount of mathematical knowledge generated so far; we communicate so slowly, in reading, writing, speaking and understanding, that it often takes years to even verify the correctness of a big result; we compute so slowly, and make so many mistakes in the process; we are emotional creatures with real-world needs, which means most of us are not willing to spend years on a hard problem with nothing to show for it. As we've discussed, we are not evolved for this, but the AIs are (or can be). The

natural conclusion, which many of us are arriving at, is that humans will no longer be at the scientific frontier pushing it forward. Science and mathematics can finally advance beyond our limitations.

Importantly, AI is here to advance our objectives. We decide what science and mathematics we'd like to see progress; we decide the problems that we want solutions to. In the past, we spent the majority of time and effort in solving our problems. Now, we should spend such time and effort on asking the right questions and understanding the solutions. For this transition, I find the following analogy appropriate: for centuries, we've been building a human library of science. We diligently built up shelves, studied and organized literature, held seminars, and most importantly, wrote new books. Now, we just discovered a massive, potentially infinite library written by another form of intelligence, in a language that we struggle to understand. It contains knowledge beyond our reach, secrets that we've long dreamt of. What is then the obvious thing to do?

Well, we should gather all of our librarians, who've done glorious work in our human library, and send them on expeditions into the library of the unknown. They will discover wondrous knowledge, bring it back, translate it, teach it, and make it a permanent part of our human library. In fact, we should significantly expand our operations, because it has never been more economical to train more librarians; the rate of return just got amplified by an immense factor. Through accessing this library, we will acquire knowledge at unprecedented speed.

There is a counterargument here, or a caveat in the analogy, regarding the human librarians' role in imparting knowledge. So far, we have imagined this mysterious library as if it is an unoccupied relic, and we are the only beings in it with agency. That is far from accurate, as in reality there will be an immense number of AI agents, way more than the number of human librarians, who would be managing this new library. They will (hopefully) be happy to assist us, to help explain things, and to help write new entries upon request. In fact, they may be able to teach or translate the material better than many of us do. Nonetheless, however capable these AI librarians are, it is always our choice what questions we'd like to ask, what knowledge we'd like to pursue. Before we make those choices, before we set such objectives,

the shelves holding such knowledge would not exist yet. Our intent and agency are crucial to the discoveries ahead.

The transition into this new dynamic has already begun. With the Lean proof of the Riemann hypothesis, which Levent Alpöge discovered together with Claude, the world of mathematics has embarked on a quest to understand the initial 17 million lines of Lean code, with the help of AI agents. We have since made significant progress. As David Hilbert, one of the greatest mathematicians in history, famously said: "If I were to awaken after having slept for a thousand years, my first question would be: Has the Riemann hypothesis been proven?" Indeed, Alpöge's proof has passed several independent Lean verifications. This is as much confidence as one can ever have in the correctness of a mathematical theorem, much more than for most results in mathematics (which have only passed human verification). It did not take 1000 years. It took 160 years of human librarian effort, around 10 years of building up the AI library, and according to Alpöge, six weeks working with the AI librarians. We will now spend months to years understanding the proof, derive new knowledge from it, and prepare new goals for expeditions into the library.

An important part of this transition is, as Terry Tao puts it, an industrialization of mathematics. Alpöge's feat would not have been possible without either Claude or an immense amount of compute. Similarly, checking and understanding the proof would take a painfully long time without computers or AI agents. For centuries, human mathematicians have been working in settings similar to artisanal workshops. We sharpen our skills, hand-craft pieces of mathematics, present them to others, and gradually accumulate peer acclaim. This century-old way of working needs to change now, much as how sewing changed after the invention of the sewing machine. The clothing industry did not disappear; the profession of cloth making did not disappear. They simply got industrialized.

The sewing workers and the Luddites of the 19th century could not, in their wildest dreams, imagine the clothing industry today. Similarly, we today may find it hard to picture the future industry of mathematics and sciences. What we do know, however, is that it will be grander than everything we've dreamt of. The human demand for knowledge, intellectual inspiration, discussions

and debate, learning and teaching, will be fulfilled to an extent greater than ever.

Us, Scientists

As idealistic as one could be about the future, we need to navigate our current circumstances, which have not been encouraging. The old system has collapsed; people are in panic or depression; lots of us chose to leave. We need to build up a new, sustainable system quickly, both for our own sanity and for the future of our craft.

This process has started across the field through seminars, discussions, and writings. I will join in with my two cents, summarized from many discussions I've had with others. Our old system rewards significant credit to priority in solving hard problems. That is not going to work any more. Going forward, we should reward

1. Asking insightful or impactful questions,
2. Good execution of research programs with AI, and
3. Communicating and teaching new discoveries effectively.

I can give an example from my field of quantum computing. At this time, there is a significant focus on building the first generations of fault-tolerant quantum computers, which will be robust against noise introduced by the real world. This is a problem where the first-order objectives are relatively clear – the target scale we want to reach is clear, and we have spent decades building up a sufficient set of methods to achieve that target. Such clear objectives mean that this is a case where AI models will showcase their incredible strength and outsmart many of us in numerous technical processes (don't get me wrong, there's still lots of human work to do in aligning the models to our numerous finer objectives). On the other hand, here is a (slightly) newer question that at least I am very interested in: can we send fault-tolerant quantum sensing devices into outer space?

Well, this single thought spins out a long line of questions. What is such a machine good for? What do we want to detect with it that we can't do on the ground? What kind of noise will our quantum devices experience, without the protection of our atmosphere but also without the ground shaking every

five minutes due to metro trains?[17] How do we engineer systems robust against such noise? Can such a device serve other purposes, such as being a medium for our future quantum Internet? I can go on and on. This is an example of breaking down a large and vague objective into many smaller and finer objectives. In this process, humans need to step in at various points to tune the objectives (e.g., what sensing task is most interesting for us?), such that the end result best serves our purposes. Anyone who can properly formulate this question, pursue it with AI, and give an excellent presentation on the findings would be a great scientist in my book. Importantly, one needs to understand the landscape of a field well enough to start asking insightful questions and setting new objectives, large or small. For this reason, it is always meaningful to learn new knowledge and existing techniques, despite the AI knowing everything there is.

With these new value functions, I believe human sciences and mathematics will thrive. The days of owning a theorem as namesake by proving it first may be behind us, but the excitement for making new discoveries will remain for as long as we continue our pursuit. For younger students who may have just started their PhD, I wish you great success in your scientific endeavor. The way we work is changing, but our sciences and mathematics, which always belong to the new generations, will be more spectacular than ever.

A few days ago, shortly after Levent Alpöge announced that he and Claude discovered a proof of the Riemann hypothesis, Ravi Vakil, president of the American Mathematical Society (AMS), made a post titled "Hurried Announcement: Alessandro Masullo GRH" (GRH stands for Generalized Riemann Hypothesis). The post states that a fellow mathematician, Alessandro Maria Masullo, has been working on a proof of GRH with their PhD advisor Brian Conrad for 12 years, and while the proof and writing still need more work, they decided to release the (unfinished) manuscript early. The manuscript is titled "The Music of the Primes". The post further states that the work is "all human, with no AI". It is now being reviewed by the mathematics community for correctness. As with most human work, there is no Lean certification, so it is fully possible that the proof may not work. Even if the overall proof strategy is correct, there will certainly be issues here and there that need to be fixed, similar to how the Poincaré conjecture was proved by Grigori Perelman, with the help and follow-up work of the whole community.

I choose to cite this story at the end of this essay not because I want to vouch for the proof's correctness, nor because I believe humans can hold a last solitary stand against AI in mathematical problem solving. I believe that battle is lost, and an incredible Renaissance is upon us. I choose to cite it because Masullo, in the Acknowledgements section of the manuscript, ended with the following:

If any message at all this work is meant to send to the mathematician who feels discouraged by the pace and tone of the current climate, it is this: we are still here. We are very much alive. We are not going anywhere.

Indeed. May we make great discoveries in the ages to come.

Acknowledgements: Many thanks to Harry Zhou, Nishad Maskara, Soonwon Choi, Anand Natarajan, Yifan Liu, and Angus Lowe for inspiring discussions.

Claude Opus 5 helped with formatting the text. The essay is written by humans and humans only, without discussions with AI models. I felt that writing this essay was a human-native task.

Notes

---

1. On OpenAI's solution to the Navier–Stokes problem, there is debate on the originality of the solution itself. We will discuss that later.↩
2. I started writing this essay in October, at which time only the first half of these events happened. Things are changing very rapidly.↩
3. See, for instance, a recent OpenAI announcement from mathematics, the history page of the OpenAI math github repo, and Anthropic's math github page.↩
4. The P vs. NP problem has a similar status, although RH has been open for longer and, for that reason, has a deeper impact on the historical development of mathematics.↩
5. For curious readers, see also the Wiki page on ZFC set theory.↩
6. The idea that mathematical laws are universal and can be used to establish communication with extraterrestrial intelligence has been used in many works of science fiction before, such as The Three-Body Problem.↩
7. Following Erdős, when we see a result with extraordinary elegance, we sometimes say that the proof is "straight from The Book".↩
8. A quote that I personally find most reflective of our folklore religion is one from The Library of Babel, which was cited in Joel Spencer's eulogy for Paul Erdős. The quote goes: "To me, it does not seem unlikely that on some shelf of the universe there lies a total book. I pray the unknown gods that some man — even if only one man, and though it have been thousands of years ago! — may have examined and read it. If honor and wisdom and happiness are not for me, let them be for others. Let heaven exist, though my place be in hell. Let me be outraged and annihilated, but may Thy enormous Library be justified, for one instant, in one being."↩
9. This is quoted from Tristan Buckmaster's statement on how he and Levent Alpöge solved a problem closely related to Navier–Stokes shortly before the OpenAI announcement. Due to the fact that they used ChatGPT extensively in their research process, there is now debate on whether the OpenAI solution to the Navier–Stokes problem is original, or heavily influenced by Buckmaster and Alpöge's unfinished work.↩
10. When these models were first trained, all training data were human generated. However, it is likely that current AIs already have other AI-

generated content in their training data.↩

11. On the TCS result, namely ETH is false, Anthropic also announced a solution by their internal model.↩
12. To be fair, modern engines still make use of a wealth of opening theories and end game theories that are derived from both human knowledge and machine knowledge. But the key evaluation function of mid-game is no longer tuned with traditional human theory.↩
13. Even $1 + 1 = 2$ needs to be further reduced down to set theory!↩
14. In an ironic way, this dream has come true in 2026.↩
15. In many ways, modern pure mathematics is like art – the value is in the heart of the people.↩
16. If you'd like a more data-based account of the sentiment in math right now (instead of my own personal account), here is a recently collected survey and some commentary: Early career researcher responses to AI in math.↩
17. An anecdote in our community is that neutral atom quantum computers at a certain lab actually get affected by a nearby metro station, because the ground shakes every so often. I'm not certain if this is true, but wouldn't be surprised if it is.↩