MATHEMATICS IN THE AGE OF MACHINES

BRYAN WANG

In summer 2024, Google DeepMind released AlphaProof, an AI system capable of producing solutions to IMO-level math problems. On one hand, perhaps the only thing surprising about AlphaProof was how soon it appeared, and not that it was possible in the first place. After all, even if only from a purely human standpoint, it is no secret that math olympiads are, more so especially in recent decades, increasingly a 'trainable' activity – it would be rare for a modern-day IMO medalist not to have spent hours in math olympiad training sessions or working on training problem sets. By its very nature, then, solving math olympiad problems should, at least in theory, be a task that is amenable to the machine learning techniques that modern AI systems are built on.

Now, if you are, say, a math olympiad student beginning to doubt the value of working on math olympiads just because of these recent developments in AI, then I would even venture as far as to say that math olympiad may not have been the right school activity for you in the first place. After all, every math olympiad problem ever has already been solved by a human, and has even been designed to be solvable by a human, within a fixed timeframe no less. Solving problems that have already been solved by humans should sound much less appealing than solving problems that could be solved by an AI.

In this sense, an apt comparison is the game of chess. It has been decades since humans lost all practical hope of ever winning (or even drawing) a game of chess against a chess engine. Modern AI-powered chess engines such as AlphaZero evaluate chess positions far better than any human chess player can. In fact, I believe computer chess games are becoming increasingly incomprehensible to even the best human chess players. Still, humans play in and watch chess tournaments, and in fact chess has experienced a recent growth in online popularity (for reasons which, to be frank, are still not entirely clear to me). Much like chess, I think humans will (and should) continue doing math olympiads, even if humans become hopelessly outperformed by AI (whatever that means in the case of math olympiads).

However, there is no denying that there are fundamental differences between chess and math olympiad. First, chess is a sport, and the human emotions which come with sport are part of what makes chess so playable and watchable. Math olympiads may be competitive in nature, but I don't think many would qualify them as sport, and in any case they are most certainly not a watchable spectator sport! In practical terms – it *may* be possible to make a living as a chess player, but certainly never as a math olympiad contestant. (By the way, humans *have* made livings off 'math contests' before – that is the story of the 16$^{\text{th}}$-century Italians, which we will come to later.)

Much deeper than that, however, is the fact that there are 'real' pieces of mathematics in math olympiad problems. What does that mean? Most of all, it raises the possibility that AI-powered systems may one day be able to tackle *all* math problems of the sort which research mathematicians are working on.

Let's examine some of the reasons why this possibility might seem like a leap in reasoning. For one, there is always this longstanding 'debate' about how accurately olympiad mathematics actually reflects research mathematics. But in my view there is no denying that there *is* mathematics being done in olympiad problems. One analogy might be the difference between chess puzzles and actual chess games. Just as chess engines do not distinguish between the two, I see no reason why a 'mathematics engine' should differentiate between pieces of mathematics coming from olympiads and research. The main difference, in my view, is simply one of scale, and experience has shown us that scale on the human level is unlikely to be a huge obstacle for computer systems to overcome.

Another key point of contention is the fact that research mathematics is, by the very definition of research, about creating new knowledge, whereas (as I have explained above) olympiad problems are, by definition, problems that have already been solved. But this overlooks the fact that olympiad problems have to be created by humans in the first place, so that the process of designing and proposing olympiad problems is, in the same vein as research mathematics, an act of creating new knowledge. In fact, one of the implicit criteria in evaluating the suitability of an olympiad problem has always been its (perceived) novelty.[1]

Much more importantly, however, I would argue that research mathematics itself is also a 'trainable' activity. Why? Well, less the teaching component, that is the entire point of a PhD program!

Now, the aim of this article is not to speculate on whether (or how soon) the possibility of AI systems outperforming humans at tackling research-level mathematics problems will arise. Nor is it to address the societal (read: economic) implications of this and related possibilities, which is perhaps a far more important issue but which lies far beyond the scope of this article. I will instead take a more theoretical approach, and focus on a much more fundamental question: what role will mathematics, and humans, play in a world where AI outperforms humans in mathematics?

When evaluating such scenarios, it is all too easy (fuelled, I guess, by our knowledge of science fiction) to fall into what I like to call the 'doomsday trap'. In an extreme form, the trap goes something like this: we begin to imagine that AI will outperform humans in any form of human achievement imaginable, and therefore humans will be rendered completely obsolete, and AI will inevitably take over the world. Clearly such arguments are neither refutable nor winnable (not to mention they involve societal implications, which I have placed outside the scope of this article), and so we should consciously avoid falling into such a 'doomsday trap' with our arguments.

[1] There is something to be said about the meaningfulness of AI-generated math olympiad problems, whenever this starts to become a thing, but I digress.

In any case, society today has in my opinion well passed the point of co-existing with technologies with the power to render our species completely obsolete (in fact, worse than completely obsolete) in one stroke. The two most obvious examples that come to mind are nuclear technologies and biological technologies. Does that mean we should stop (or tightly regulate to the point of being impossible) all research on nuclear and biological technologies? Or should we start to view all nuclear and biological technologies with cynicism or antagonism? No! In fact, continuing research in nuclear science and biological science over the last half-century has led to some of the most promising advances in energy and medicine, for the benefit of humankind (if used correctly).

Of course, this is just to say that not all is lost even if we accept some premises of the 'doomsday argument'. We all know that AI can be a force for good, an example being the work of AlphaFold on protein folding. But the astute reader will already have noticed that there is something fundamentally different at play when it comes to the scenario of AI doing mathematics. The crux of the issue is none other than the direct replacement of human effort. In short, no human has spent their career trying to produce nuclear-levels of energy by hand. But plenty of humans have spent their careers proving mathematical theorems by hand, and if AI can outperform humans in this task, then we are led directly to the fundamental question which I want to address in this article.

It is at this point that I should make absolutely clear: just because AI can perform certain tasks, does not in any way diminish previous (or current) human achievement in these tasks. Already I have stressed above the two examples of chess and math olympiad. But even within 'mathematics proper', history is replete with examples of human mathematical tasks being completely superseded by computational advances. Mathematicians are probably all familiar with the story of how 16^{th}-century Italian mathematicians would challenge one another to solve polynomial equations, many of whom devoted their careers to (and indeed staked their careers on) such tasks. But this is of course a mathematical task which would be considered completely mundane today, because we now have computational tools like WolframAlpha that can solve polynomial equations in the blink of an eye. Yet it would be ridiculous to claim that the Italians have wasted their efforts trying to solve polynomial equations – and this is not even to mention that their efforts led them directly to the notion of *complex number*, which, as we all now know, went on to shape mathematics in unimaginable (no pun intended) ways.

Some will argue that this is an unfair comparison as the Italians lived 500 years ago. But in my view this is forced upon us by the exponential nature of technological advancement – I suspect that climate change campaigners will find this a depressingly familiar challenge.[2] We are fortunate, however, that mathematics has the almost-unique luxury of being a field old enough that we are able to make such comparisons

[2]Some might also argue that this is an unfair comparison as the development of modern computational tools was only (indirectly) made possible by the work of the Italians. But there is an important subtlety here: virtually all our current AI systems gain their abilities by learning from (massive amounts of) data, which in our case would almost certainly include the work of many human mathematicians,

without resorting to hypotheticals. And what we see is that while much of the work of the great mathematicians of centuries past is now considered completely mundane in the eyes of modern research mathematicians (think the development of calculus, but there are of course many other examples), this work is still equally celebrated as pinnacles of human achievement.[3] And even today, mathematics students around the world are routinely tortured with differentials and integrals (among others), when the vast majority of which would take WolframAlpha mere seconds to perform.[4]

From a purely theoretical and mathematical perspective, therefore, I believe the crux of the issue is not so much the replacement of human effort *per se*, but the fact that (a) AI systems today are increasingly capable of carrying out mathematical proof, and (b) our present mathematical culture is centred around the provision of proofs.

This latter statement deserves explanation, which I will come to shortly. But first let me say that mathematics *itself* is also concerned with the study of mathematical proof – that is the field of mathematics known as mathematical logic – and this study is exactly what makes it possible for computer systems to produce mathematical proofs in the first place (via the formalization of mathematical proofs). And the idea that the production of, or search for, proofs, can be automated – in other words, performed by an algorithm – is certainly not new. In fact, this idea goes back even to the very origins of the field of "artificial intelligence". In 1956, Allen Newell, Herbert Simon and Cliff Shaw wrote the "Logic Theory Machine", a computer program that has since been described as "the first artificial intelligence program". Simon would later recall of their efforts: "In the fall of '55 we decided that a chess machine was not the thing to start on – that an easier task was to build a theorem prover."(!) [Ma].

I would like to propose, therefore, that in a world where AI systems outperform humans at providing mathematical proof, proof will come to be seen as a form of computation. In other words, we will view proofs as we view the results of computations today.

For now, let us restrict our imagination to a world where AI systems outperform humans at providing mathematical proof, and mathematical proof only. Part of the reason is of course to avoid the doomsday trap I have mentioned – if we begin to imagine that AI systems will perform any and every other task conceivable by a human, we will get nowhere, and so we have to first draw the line somewhere concrete. But more concretely, perhaps now is a good time to confront the fact that all our current AI systems are, from a purely theoretical standpoint, nothing but algorithms running on a computer, which, again, mathematics itself has provided us a way to understand.[5]

past and present. So the work of human mathematicians may in fact contribute in a much more direct way to the development of any future AI system, than one might first expect.

[3] Sometimes I get the feeling that some non-mathematicians actually admire these achievements more than mathematicians ourselves do – a classic case of the curse of knowledge, perhaps?

[4] Needless to say, a whole other article could be written about issues in mathematics education, of which I will make no comment in this article.

[5] I have shamelessly claimed theoretical computer science as a branch of mathematics here, and I hope that not too many will disagree. Certainly this is undeniable from the historical standpoint – mathematicians reading this probably need no introduction to Turing's eponymous machines and his

Therefore it is only fair that we restrict our imagination to tasks which we already know are possible, in theory, to be performed by an algorithm.

The point, of course, is that even this restriction does not seem like much of a restriction when it comes to the broader notion of AI outperforming humans in mathematics *per se*. The reason, in my view, is only because of our present mathematical culture. Just as the mathematical culture of the 16$^{\text{th}}$-century Italians centred around providing solutions to polynomial equations, I would like to say that our present-day mathematical culture is centred around providing proofs to theorems. Now obviously this is not to say that this is the *sole* goal of mathematics today, nor in the 16$^{\text{th}}$ century – the Italians were no doubt interested in other things beyond solving equations (I have already mentioned how they were the ones who developed the notion of complex numbers). Rather, this is a reflection of what we *value* the most in mathematics – proofs for us, and solutions to equations for the 16$^{\text{th}}$-century Italians.[6] If we were to travel back in time and bring WolframAlpha to the Italians, it would probably have caused a seismic shock among the mathematicians of the time – indeed, many of them might have considered such computer systems to herald the end of mathematics as they knew it. But today we would hardly consider the computations performed by WolframAlpha to be groundbreaking in any way. (It is worth noting also that very few (pure) mathematicians today actually understand all the details of how WolframAlpha actually works.)

Now some might argue that this is a far-fetched example, and I must concede that it is not a perfect comparison in some respects. But the point I want to make is this: proof is just a *part* of mathematics, which happens to be the part we value the most right now.

This raises the obvious question: if not proof, then what else makes mathematics *mathematics*? It is difficult to imagine, but that is precisely only because of our current mathematical culture which is so deeply centered around proof. Some will have suggested that mathematicians today deal in other intangible things besides proof: intuition, visualization, and analogies, just to name a few. But this does not change the fact that everything still centers around the provision of proof. We develop intuition in the hope of finding new proofs. We develop analogies in the hope of finding new theorems to prove. Besides, things like intuition and analogies are most certainly not unique to mathematics as a field (and whether they are even unique to humans is a completely different rabbit hole which I will not go down for the moment – recall the doomsday trap), and so are not suitable defining characteristics of mathematics.

It is at this point that I would like to propose a working definition of mathematics in a 'post-proof' world, which you could consider the central proposal of this article.

---

solution to the decision problem (which asks, in layman's terms, if every mathematical question can be answered by a computer program. Spoiler alert: the answer is no.).

[6]Cardano, who conceived of complex numbers in his monumental work on algebra, *Ars Magna*, famously described them as being "as subtle as they are useless"! This was, of course, because the Italians only cared about the real solutions to their equations (pun intended). There are many more examples throughout history of mathematical discoveries being severely underestimated even by their discoverers, and I will leave the joy of discovering more such examples to the interested reader.

Mathematics is the human process of building and understanding *human models* of not only what exists (the Universe), but also what could exist (future technologies), and what may never exist (human ideas and concepts). These models must be precise and reproducible, but more than just that – **they are human mental models with unexpected and unimaginable applicability across a wide variety of domains**, and that is what gives mathematics its value, appeal, utility and importance.

Now such a description may seem hopelessly vague – but really, which 'definition' of mathematics doesn't suffer from the same problem? Let me try to dispel some of the likely objections I anticipate from mathematicians (and non-mathematicians). First, the word "model" is loaded, perhaps even more so to mathematicians, many of whom have come to associate it with the "mathematical modelling" of applied mathematics (more on the distinction between 'pure' and 'applied' mathematics below). That is why I have specified that these models should have unexpected applicability. The whole point of "mathematical modelling" in the current sense is to model a specific, *given*, situation or problem, and so by definition there should be nothing unexpected about their applicability. Furthermore, if a completely different situation or problem is requested to be modelled, the specific models already used for previous situations should not apply – otherwise the entire field of "mathematical modelling" would scarcely exist! Perhaps another possible word to use in place of 'model', if you'd like, is the word 'structure', but that seems to be a loaded word as well. (If we already had the perfect word to describe mathematics, then this would not be such a difficult question in the first place!)

Second, you might like to think that the reason for this wide applicability is solely due to abstraction. I would like to deliberately avoid using the notion of 'abstraction' to define mathematics, for three reasons: first, it is not at all clear (and highly context-dependent) what the word 'abstract' means; second, many mathematical objects are rather concrete objects to mathematicians; and third, it is not the pursuit of abstraction that leads to the wide applicability. Mathematics has never really been about the pursuit of abstraction for abstraction's sake – that in my view falls more in the realm of philosophy (incidentally, abstraction and philosophy are exactly what we are doing in this article, and I am sure you will agree that this article is not mathematics).

Finally, the word "applicability" is also an especially loaded one to mathematicians, due in no small part to the modern distinction between 'pure' and 'applied' mathematics. I am of the (not too controversial, I hope) view that unlike, say, the distinction between organic and inorganic chemistry, which is an inherent distinction dictated by the chemistry itself, the distinction between pure and applied mathematics is largely an artificial one (by which I mean: largely driven by societal considerations), and not an inherently mathematical one. The sole difference, to me, is that of intent: 'applied' mathematicians study these *human mathematical models* with their applications in mind, while 'pure' mathematicians study them for their own sake. But at the end of the day, the outputs of 'applied' mathematics *are* mathematics, no less. And if the history of mathematics has had anything to say, it is that intent and outcome are completely separate things – we can be certain neither Newton nor Leibniz had in mind any of the fantastical modern applications of calculus, but of course calculus is today the backbone of almost all 'applied' mathematics. (As an oft-cited example, the well-known backpropagation of machine learning is in my view really a glorified version of Leibniz's chain rule.)

Incidentally, it must be said that Newton and Leibniz's version of calculus falls well short of the standards of mathematical rigour (let alone proof) of our modern mathematical culture – yet clearly there is something we value greatly about their version of calculus (or we really wouldn't care if Newton or Leibniz should get the credit). So the invention of calculus is arguably the simplest and most famous example of something, beyond proof, that we value in mathematics even today.

There is an analogy that I especially like when it comes to the symbiotic relationship between 'pure' and 'applied' mathematics. The 'pure' mathematician is like a train track builder – he just enjoys laying train tracks for their own sake, with not much regard for who might want to use his train tracks. One day the 'applied' mathematician wants to get somewhere, and finds that, what do you know, there are already some train tracks there leading him right where he wants to go! Though their intentions are very different, both 'pure' and 'applied' mathematician are contributing to one and the same human enterprise, which we call the railroad industry – there is no "train track industry" without the "train carriage industry", and vice versa.

That is where the analogy usually ends, but for the purposes of this article we can take the analogy one step further. One day, a machine is invented to automate the laying of train tracks, far faster and further than any human being can. There is much fanfare, anticipation, (and handwringing about it by the 'pure' mathematicians). The next day, the machine is deployed, and in a matter of days every corner of the Earth is absolutely covered in train tracks, with rows and rows of tracks criss-crossing one another on a completely human-incomprehensible scale. It doesn't take much imagination to realise that an Earth completely tiled by train tracks is no better than an Earth with no train tracks at all. (If you'd like, you can also replace train tracks by asphalt roads, and train carriages by cars, in this analogy.) Clearly there is an art to the building of train tracks – whether it be building tracks that lead places where humans consider beautiful, or making sure that the tracks form human-coherent lines and networks – which cannot simply be replaced by a superhuman track-laying machine.[7]

I now want to focus on the singular most important word in my description of mathematics, and that is the word 'human'. I have always held the belief that **mathematics does not exist without humans**.[8] Well, then what exists without humans? Certainly, the Universe exists without humans. Some may then say that the Universe operates according to certain mathematical laws. I am of the view that, no, it does not! The Universe operates as the Universe operates, and these mathematical laws are precisely our *human models* of the Universe. And none of these human models actually describe the *actual* Universe: they may describe isolated things from our Universe in

[7]By the way, to imagine that the machine can start inventing and building alternative forms of transportation, like planes, would be to fall into the doomsday trap. In any case, the same can be said if the machine starts tiling the Earth with airports and runways.

[8]I am aware, as with all things in this arena, that there are longstanding philosophical debates around these issues.

a closed system, without any external effects, and so on and so forth. But they allow for a *human understanding* of our Universe, and it so happens, completely and utterly surprisingly I should say, that these models are written in the language of what we humans call mathematics. Some of these models even lack the sort of rigorous foundations that today's mathematical culture demands, yet humans have nonetheless developed ways of understanding and working with them, and they have generated quantitative predictions which agree with experiment to, frankly, mind-blowing accuracy.

The same goes for computer-produced proofs. I have already said that the computer production of proofs has thus far only been made possible by the formalization of mathematical proof, itself arising from the branch of mathematics known as mathematical logic. A formal proof, in the absence of humans, is nothing more than a collection of symbols; but these collections of symbols are understood by humans via our *human mathematical model* of mathematical proof. I have also said that all current AI systems are nothing more than algorithms running on computers (and that to go any further would again be to fall into the doomsday trap), which we understand via our *human mathematical model* of computing machines, à la Turing (and others). In the absence of humans, the output of these computer algorithms is *not* mathematical proof, but really nothing more than a series of 0s and 1s. Actually, the very notion of 0s and 1s is also a *human mathematical model*! What really exists, without humans, is the lack of or presence of electrons flowing through whatever electronic components were used to build the computer system, which humans then understand as a 0 or a 1.[9] The same can be said of any future computer system, such as the quantum systems commonly touted in recent years. The system is as good as non-existent without human interpretation, which is only made possible by our human, mathematical, models of it.

If there is one, and only one, article that I think every mathematician should read, it is "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" by Eugene Wigner. The first paragraph reads as follows:

There is a story about two friends, who were classmates in high school, talking about their jobs. One of them became a statistician and was working on population trends. He showed a reprint to his former classmate. The reprint started, as usual, with the Gaussian distribution and the statistician explained to his former classmate the meaning of the symbols for the actual population, for the average population, and so on. His classmate was a bit incredulous and was not quite sure whether the statistician was pulling his leg. "How can you know that?" was his query. "And what is this symbol here?" "Oh," said the statistician, "this is pi." "What is that?" "The ratio of the circumference of the circle to its diameter." "Well, now you are pushing your joke too

[9] Binary numbers were first employed in the study of mathematical logic by none other than Leibniz, who has been described by some as the "first computer scientist". His theoretical *calculus ratiocinator* (and physical "stepped reckoner", among others) are important precursors to our modern-day notion of computing and reasoning machines. As described in [Jo], Leibniz envisioned "*a calculus ratiocinator in which the rules of reasoning are translated by laws like those of algebra, and reasoning becomes a machinelike calculating process which frees the imagination where its action is not essential and thus increases the power of the mind.*"

far," said the classmate, "surely the population has nothing to do with the circumference of the circle."

I think this simple example already serves as the most poignant illustration of what I mean by the "unexpected and unimaginable applicability" of our human mathematical models like circles and pi. Similarly, why should the ratio between two sides of a right triangle have anything to do with waves, or signal processing (via Fourier analysis)? I will leave it to you to come up with more examples, of which I am sure any mathematician can easily find many.[10]

On the second page of the article, Wigner directly addresses the question, which, if it is not already clear by now, is really the central question we should have been asking ourselves all along: "What *is* mathematics?"

He gives the following description:

" [...] mathematics is the science of skillful operations with concepts and rules invented just for this purpose. ***The principal emphasis is on the invention of concepts. Mathematics would soon run out of interesting theorems if these had to be formulated in terms of the concepts which already appear in the axioms. "***

If there is one thing I hope you take away from this article, it is that mathematics is *not* just proof, nor is it just the act of proving theorems. And even if we find it still difficult to accept a vision of mathematics where proof is not central, we cannot escape the fact that the proof-centered mathematics culture of today must also be concerned with *what* to prove, and *why* we prove what we prove. I think we will find, as mathematicians, that there is still much to be answered about these questions.

I want to end this article by conducting a little thought experiment. Suppose we allow ourselves to fall ever so slightly into the doomsday trap, by imagining an AI system that outperforms humans in any and all thinking tasks, even tasks like the generation of ideas. This means, for instance, that AI will be better at designing new computer systems, and hence by extension new AI systems, than any human can. One can quickly see how this can again easily spiral us down the AI-doomsday trap, in true sci-fi fashion, and there is really nothing much to say in this regard.

But putting this aside, it would also mean that we imagine an AI with a better 'mental' (mathematical) model of the Universe than any human. This model need not describe the Universe perfectly (indeed, it obviously cannot), it just needs to achieve a more complete understanding than any human being can. By assumption, this model should basically be a complete black box to even the best and brightest humans (indeed, large parts of many of our current AI systems already behave essentially as black boxes to us, or we would not call them AI). Even if this AI system were capable of explaining itself to humans – given that it outperforms humans in any thinking task, which includes the explanation of concepts to humans – we might as well assume (since, again, it

[10] I meant this article also for a more general audience, so I have throughout deliberately avoided examples which require more 'specialized' mathematical knowledge. I leave it as an exercise to the reader to find more 'sophisticated' examples of many of the points I raise in this article.

outperforms humans in every way) that the explanations would be incomprehensible to any of us *within our lifetimes*, and so as good as useless.[11]

This is an important point, and so it helps to have some concrete illustrations in mind. Continuing the analogy with chess, imagine how incomprehensible computer chess games already are to the best human chess players, with calculations running tens of moves deep which would be incomprehensible to humans within our lifetimes – perhaps chess engines have an 'understanding', whatever that means, of certain advanced chess principles or heuristics, but which would be completely unexplainable to humans, let alone implemented by humans in practice. Or, imagine this as something like giving the ancient Egyptians a human oracle with the knowledge of all of modern science (but not history, since that would be knowledge of the future) – how would the Egyptians even know what to ask her, and where would she even start in getting the Egyptians to produce and implement, let alone understand, modern technology within their lifetimes?

What could a human-comprehensible output of such an AI system be, then? Well, in the first place, the other core output of fundamental science, besides human understanding, is to make predictions about the Universe (recall reproducibility is one of the key principles of science). So, the conclusion is that such an AI system would be a black box, yield no human-comprehensible insight or understanding, but does give us predictions about the Universe (assuming we understand them). We already have such a thing – it's the damn Universe itself! Running this AI system would just be like running the damn experiment, and it's not even clear if it would take far less resources to run the AI – already, one of the problems we are beginning to see with our new AI systems is the sheer amount of energy and resources they require. *If this AI suffices for our needs, then perhaps we never had any need for theorists in the first place.*

A final saving grace for this AI system, some might say, is that we don't even have to bother about human comprehensibility, or human understanding. It can just straight up tell us what to do to advance science and technology (some fancy terms can be applied to this notion, such as calling this the replacement of human creativity or human innovation). But chess engines have, since their inception, been telling us what to do, by telling us the best next move, and we are yet to see any evidence of humans demonstrating an understanding of chess at the level of chess engines, and indeed, evidence of such is nowadays almost always taken by the chess community as evidence of cheating. And in the case of the Egyptians? Even if they believe in the benevolence and correctness of the oracle (whatever that means), having the oracle tell them what to do at each step would not only *not* give the Egyptians any more understanding of modern science, it would herald the end of Egyptian civilization (actually, if the oracle forces matters, the more likely outcome would be the lynching of the oracle by the Egyptians). You can easily imagine that anyone stuck with an ancient-Egyptian-level of knowledge, living in a modern society, with no way to understand modern knowledge, would be completely dominated by modern humans, and in this case there would be

[11]Here we need to make the assumption that this AI system has not yet managed to implement any sort of bioengineering of humans to enhance our thinking capabilities or lifespans, or we would well and truly have fallen into the sci-fi-fuelled doomsday trap.

no modern humans to speak of – they would be completely dominated by the oracle. My point is that if humans begin to do things which no human even *believes* they understand, then we must really begin to question if the foundations of our civilization have not already been radically altered or replaced with something completely different. In other words, this is again the doomsday trap in disguise.[12]

In the ideal (and in my view most likely) scenario, just as humans have come to use chess engines as an aid to help analyse their games, the Egyptians would come to use the oracle as a guide, consulting her on issues they face or things they don't understand. Our understanding of chess will accelerate, as will the technological progress of the Egyptians, no doubt. But humans will have to understand chess in our own terms in order to play good chess, just as the Egyptians will have to come to develop an understanding of science and technology in their own terms, in order for them to meaningfully use it. Even in an age of superhuman AI, humans will have to understand science and technology in our own terms, in order to do anything meaningful with it. And it just so happens that our human mental models of science and technology are written in a language we call mathematics.

Speaking of language, let's at last return back to present-day reality, and remind ourselves of the development which arguably sparked all the events that led up to the writing of this article: the rise of Large Language Models (LLMs). It is undeniable that LLMs have today mastered language at the human-level, and occasionally the superhuman-level. I believe that, as with many other technologies, the development of LLMs will on balance do more good than harm for the human race.[13] But only a lunatic would suggest that we should stop teaching language to our children, or stop teaching all but the simplest of words ("eat", "drink", "sleep"), and let LLMs fill in all the rest for us. Why not? After all, humans can and will survive just fine without the intricacies of human language – just ask literally any other living animal on the planet. And not every human has to appreciate, nor be interested in, nor want to study at a deeper level things like literature, or poetry, or linguistics, or the science of human communication. But we as humans have come to understand that having more than a basic, survival-level, understanding of language is part of what makes us human (and not animal), and that to lose this understanding would be to lose a part of what it means to be human. I have every belief that we will come to the same understanding about mathematics.

References

[Ma] D. MacKenzie, *The Automation of Proof: A Historical and Sociological Exploration*, IEEE Annals of the History of Computing, Vol. 17, No. 3, 1995.

[Jo] P. E. B. Jourdain, *The Logical Work of Leibniz*, The Monist, Vol. 26, No. 4 (Oct 1916), pp. 504-523.

Email address: `bwangpengjun@math.harvard.edu`

[12] I must emphasise again that in this extreme (but not implausible) scenario, the design and maintenance of AI systems will also be a task impossible for humans to perform, at least not without significant AI assistance, since it is also a purely thinking task.

[13] The closest example is that of search engines, which has unquestionably done more good than harm, and most certainly not led to the obsolescence of things like physical libraries.