Hester Breman
Faculty of Psychology and Neuroscience
Maastricht University
h.breman@maastrichtuniversity.nl

Event Workshop Lorentzcenter, Leiden, The Netherlands, 15–19 September 2025

Mechanization and mathematical research

While Ada Lovelace already envisioned that machines could do complex mathematics in 1843, medic McCulloch and self-taught logician Pitts conceptualized in 1943 a machine that could actually learn, by modelling how neurons in the brain worked using logic. Psychologist Rosenblatt built a learning machine in 1958 [1]. And here we are in 2025, with companies and computer scientists racing to develop artificial intelligence rivalling mathematicians [2]. This development led to the workshop “Mechanization and mathematical research” from 15–19 September 2025 in the Lorentzcenter, Leiden, where questions like “Why do we do mathematics?”, “Where has the play gone?” and “Will mathematicians die with a hammer — or piece of chalk — like John Henry, the steel driving man?” were discussed. With participants from the fields of mathematics, politics, philosophy of mathematics, cognitive neuroscience, history and artificial intelligence (AI), from PhD students to Fields Medal winners, from luddite to hardcore formalizer, this momentous workshop hosted interesting and essential discussions and the organizers made sure the workshop would be followed-up by actions.

History

The idea for the workshop started by collaboration between Lenny Taelman (University of Amsterdam), Rodrigo Ochigame (Leiden University), Mateja Jamnik (University of Cambridge and the Leverhulm Centre for the Future of Intelligence; see the interview on page 208 in this issue), Akshay Venkatesh (Institute for Advanced Study (IAS), Princeton) and Johan Commelin (Utrecht University). The goal was to bring together key figures in the field and interested participants to reflect on the future of mathematics, considering that machines already can solve most problems from international math olympiads [3], help generating conjectures [4][5][6] and prove theorems [7]. Even though manually programmed symbol manipulation machines performed mathematics before, for example The Logic Theorist in 1956, Automath in 1967 [8] and the HR program in 2000 [9], neural networks that can actually learn are since 2012 successful, with one of the reasons more data [4,16]. Later, techniques such as transformers improved achievements of neural networks even more.[10]

Day 1

After a welcome by Ochigame, Emily Riehl explained how computer proof assistants or interactive theorem provers (ITPs), such as Agda, HOL Light, Isabelle, Lean, Mizar, or Rocq (née Coq) can help writing formal proofs. ITPs know the rules of a logical formal system (e.g., a foundation for mathematics) and a trusted core program (the kernel) uses the rules to check the correctness of proofs. While in the past, the foundation was based on traditional Zermelo-Fraenkel set theory and first order logic, modern proof assistants use a newer formal system: type theory,[11] where logical propositions and data types are “types”. Asking what the point would be of explaining proofs to a computer via formalization, Riehl identified ‘rigor’, ‘clarity’ and ‘gamification of mathematical research’. Foreseeing how AI-powered autoformalization will drastically change every day practice of mathematics, she concluded with the question “How would this shift affect human understanding of mathematics— the real point of what we do?” [12]

Michael Barany (University of Edinburgh) continued the narrative by pointing out how mathematical practice is embedded in society with political-economical dimensions and physical reality (“stuffness”).

Albert Jiang gave his view on developments in AI for theorem proving, briefly including his work on a method that maps informal proofs to formal proof sketches.

Day 2

The second day started with a lecture by Walter Dean (The University of Warwick) on the potential role of computational complexity in the philosophy of mathematics. He stated that technological advances do not help with intrinsically hard problems.

Kevin Buzzard (Imperial College) reported on how he involved students

Participants of the workshop "Mechanization and mathematical research".

and postdocs in formalization; the latest project by postdocs is writing the statements of theorems published in *Annals of Mathematics* in Lean code. He sketched the possiblity of a dystopian future where machines prove the Riemann hypothesis all by themselves, at which journal editors responded that the future is already here: they are flooded by papers proving the Riemann hypothesis aided by large language models (LLMs). Buzzard's dream is that machines would verify what we know, discover what we don't know and help regain interest in curiosity–driven research and concluded with "...until machines start doing that, too".

Catarina Dutilh Novaes (VU Amsterdam) characterised proofs as a dialogue. She then continued wondering whether computational devices are, like diagrams, also joint epistemic actions, for example in a dialogue between a mathematician and an LLM.

Dame Ursula Martin (University of Edinburgh) led a panel discussion featuring Novaes, Dean, Sophie Morel (ENS de Lyon) and Buzzard, about whether mechanisation is changing our idea of what a proof is; also, how errors sometimes are propagated in journals, considered problematic by some and fruitful by others. Then, Fabian Gloeckle (École nationale de ponts et chaussées) explained methods used to scale inference compute in theorem provers. One is tree search such as Monte Carlo Tree Search (MCTS), which has the disadvantage that it does not have a global state. Another is reinforcement learning (RL) — a method based on rewards and penalties — which can work out well in case limited data is available.

On the theme "Proof", participants assigned themselves to one of the breakout rooms of their interest.

Day 3

On the third day, Heather MacBeth (Imperial College) started by elucidating how computation in mathematics has been valued differently throughout history.

Simon DeDeo (Carnegie Mellon University, Santa Fe Institute) entertained the participants with a SAT puzzle as an introduction to the role of causes in mathematics. According to him, these are in mathematics not causes of the facts, as in science, but causes of a mathematician's knowledge of the facts.

Robert Y. Lewis (Brown University) advocated the value of formal mathematics as subdiscipline of mathematics, noticing a vibrant community, intensively communicating via for example Zulip and GitHub.

The lectures were followed by a discussion about "What is a mathematician? What do mathematicians do?" between DeDeo, MacBeth, Lewis and Barany chaired by Antoine Chambert-Loir (Université Paris Cité).

Jeremy Avigad (Carnegie Mellon University) introduced the Institute for Computer-Aided Reasoning in Mathematics. He observed the declining use of logic in artificial intelligence (AI) and asked the question: "Is mathematics obsolete?", accompanied by music about the Americal folk story about John Henry, the steel driving man, who lost his life due to a competition with steel drills.

The gloomy song was followed by discussions in breakout rooms, including the topics "protein shakes ('Is math hard?')", intuition, inclusion & safety in the Zulip community, and teaching.

Day 4

Talia Ringer (University of Illinois at Urbana-Champaign) gave a glowing speech with the title "Big Pretty Math", arguing how proof assistants can help bring fields and people together, with the example of Kim Morrison suddenly collaborating with Terence Tao as a result of a "pull request" (a procedure in online software projects requesting to include new code).

Andreas Stylianides (University of Cambridge), foreseeing machines doing the proofs, held a heartfelt plea for changing the educational curriculum by emphasizing creativity, the role of intiuition and proof construction.

Massimiliano Gubinelli (University of Oxford) gave an engaging talk on the role of human stories in mathematics in the age of machines.

Duing a panel session chaired by Bryna Kra (Northwestern University), Slava Gerovitch (MIT), Gubbini, Styliniades and Ringer considered the question what the purpose of mathematics is, and "why do we do mathematics?" on which the answer was "because we like it".[13]

Michael Harris (Columbia University) stated during his talk with the title "Is mathematics obsolete? Why do we ask why do we do mathematics?" that the business model of Siilicon Valley is based on monetizing human experience. Referring to his blog "Silicon Reckoner"[14], Harris described why according to him math is from Eros: alive, and and art form, and computing is from Thanatos: frozen into 0s and 1s. Harris quoted Melanie Mitchell (Santa Fe Institute) in how AI is a double edged sword, and Kate Crawford who observed how AI depletes resources such as water and energy [15]. Then, Harris generously shared with all participants some home–made quince paste, asking “If you can buy it in the store, is there still a point of making it?”

Day 5: Will mathematics still exist in 2035?
In the morning Fields Medal winner Timothy Gowers (Collège de France, University of Cambridge) wondered whether mathematics will still exist in 2035. Gowers described nine possibilities for the next ten years and estimated the likelihood of each: for example the chance of an external disaster (e.g. nuclear war) low but not zero and the chance of gradual takeover (e.g. combinatorics might be ruled by machines, but algebraic geometers still independent), Gowers estimated high.

The final breakout session resulted in recommendations to distinguish between the interest of mathematicians and corporate interests, emphasize open source software, for conscious agenda setting with politicians, to include the Global South and to update software licenses. Also, a committee was formed to write ”The Leiden Declaration of AI and Mathematics” coordinated by Jim Portegies (Eindhoven University of Technology), Ochigame and Harris.

Public Symposium “Mathematics in the age of automated proofs”
Robbert Dijkgraaf (former director of the IAS in Princeton and Minister of Education in the Netherlands) sketched future scenarios, which could range from utopian, dystopian to business as usual. He was certain that the impact of the workshop will be felt. He then introduced the speakers Thomas Hubert (Google DeepMind), Stephanie Dick (SFU Vancouver) and Akshay Venkatesh (Institute for Advanced Study (IAS), Princeton).


Symposium “Mathematics in the age of automated proofs” with Dijkgraaf (chair) and speakers Hubert, Dick and Venkatesh.

Hubert explained that AlphaZero — a system that learnt to play chess and Go from scratch and achieved super-human level — and AlphaProof, which interacts with the Lean proof assistant, work so well due to reinforcement learning (RL), a technique where systems learn by themselves (“unsupervised”). He called SuperScale RL a proven recipe to superintelligence. Another AI mathematics system developed by DeepMind is AlphaTensor, which discovers new algorithms, such as matrix multiplication.

Dick pointed out how the occupation of mathematicians has changed over centuries. She illustrated how the Logic Theory Machine (1956) by Newell & Simon at the RAND corporation used abstraction because of uncertainty in future geographical, political and military factors, while philosopher of mathematics Hao Wang thought computers should be part of the messy reality. Dick envisioned the occupation of mathematicians is likely to change in future, but probably will continue to exist.

Venkatesh reported about a letter he received from a student containing the question: “Is AI destroying the soul of what we find interesting?” To answer this question, he focused on the experience of mathematics and the intrinsic value of mathematics. An extrinsic value would be how useful mathematics is for other sciences; as intrinsic values Venkatesh identified mathematics as an exact language for thought and communication, and mathematics as a form of play.

To conclude, Dijkgraaf asked the panel members for a positive message for the future of mathematics. With Hubert endorsing the following views, Venkatesh asserted that mathematics is an essential part of what it means to be human. Dick agreed and indicated she would love for mathematics to return to older philosophical questions, which AI might actually make possible.

Conclusion
This historical workshop clarified the role of narrative, emotions and play in human mathematics, provided a clear overview of the current state of mechanization in mathematics and showed that the way forward for mathematicians is to unite, connect with politics and educate students considering future technology. Finally, the discussions served as an appeal to get involved to keep mathematics humane.

Notes and references

1 G. Lindsay. *Models of the mind; how physics, engineering and mathematics have shaped our understanding of the brain*. Bloomsbury Sigma, 2021.

2 M.N. Rabe and C. Szegedy. Towards the automatic mathematician. In *28th International Conference on Automated Deduction*, volume 12699 of LNAI, 2021.

3 https://deepmind.google/discover/blog/

4 Y.H. He. Machine-learning mathematical structures. *arXiv*, (2101.06317), 2021.

5 M. Johansson and N. Smallbone. Exploring mathematical conjecturing with large language models. In *Proceedings of the 17th International Workshop on Neural-Symbolic Learning and Reasoning*, 2023.

6 A. Davies et al. Advancing mathematics by guiding human intuition with AI. *Nature*, 600, 2021.

7 K. Buzzard. Proving theorems with computers. *Notices of the American Mathematical Society*, 67(11):1, 2020.

8 J. Harrison, J. Urban and F. Wiedijk. History of Interactive Theorem Proving. In: *Handbook of the History of Logic* 9: 135-214, 2014.

9 S. Colton, A. Bundy, and T. Walsh. Automatic invention of integer sequences. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on on Innovative Applications of Artificial Intelligence*, 558–563, 2000.

10 See also the “New Technologies in Mathematics” seminar at Harvard CMSA: https://www.youtube.com/playlist?list=PLoNRmB-ofnLJQJ3fuIk3yVULtm6_JnQ_zI.

11 See for example E. Rijke and B. Spitters. Homotopy type theory and the formalization of mathematics. *Nieuw Archief voor Wiskunde 17(3)*, 2016.

12 https://emilyriehl.github.io/files/proof-paradigm-lorentz.pdf

13 Which is confirmed by research on motivation for research mathematics: 56% of respondents chose ‘direct emotional satisfaction connected to the activity’; see H. Breman, R. Hoekzema, H.K. Sørensen, M.W. Johansen and R. Goebel. On the importance of visuo-spatial thinking for research mathematicians. *Canadian Journal of Experimental Psychology*, 2025.

14 https://siliconreckoner.substack.com/

15 K. Crawford. Atlas of AI: *Power, Politics, and the Platetary Costs of Artificial Intelligence*. Yale University Press, 2021.

16 J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, and F.F. Li. Imagenet: A large-scale hierarchical image database. In *Conference on Computer Vision and Pattern Recognition*, 2009.