Automated Theorem Provers and the instruction of proof-based mathematics

Dr. Ben Davies, West Virginia University

A degree in mathematics is often referred to as an apprenticeship into the community of mathematicians (Dawkins & Weber, 2017). Yet, by 2040, it is possible that many students may never physically meet a mathematician or even set foot in the ivory towers in which mathematicians conduct their business. In the context of a global pandemic, the shift toward virtual instruction has rapidly accelerated. But make no mistake, this is an acceleration, not a change of direction. During 2020, many institutions moved online in a matter of weeks, but the shift toward virtual delivery of higher education has been gradually accelerating for more than a decade.

Massive Open Online Courses (MOOCs) for introductory calculus, linear algebra, geometry, and combinatorics are becoming increasingly popular. Yet, more advanced aspects of the tertiary curriculum have remained fundamentally unchanged. Instruction of proof-based mathematics has resisted the shift toward virtual instruction and is often viewed as a gatekeeper to upper-division mathematics. In 2020, there are no fully autonomous courses in proof-based mathematics, meaning that those seeking a mathematics degree must have at least human-to-human interaction to progress toward more sophisticated mathematical ideas.

This status quo will not survive recent technological developments. With recent advancements in the pedagogic value of Automated Theorem Provers (Avigad, 2019), and in computer-aided assessment protocols (Sangwin, 2013), a fully autonomous introduction-to-proof course is increasingly possible. This has substantive consequences for tertiary mathematics education. Here, I focus on the well-studied 'introduction to proof' setting but note the potential for similar developments to permeate throughout the tertiary curriculum, unlocking the prospect of complete, fully automated degrees in pure and applied mathematics.

MATH2XX: A fully autonomous introduction to mathematical proof

This (imagined) second-year course features extensive use of ATPs, fully automated assessments of students' understanding, and personalized feedback for each student.

Automated Theorem Provers, such as LEAN and Coq, have seen rare but increasing usage in undergraduate instruction on proof (Avigad, 2019; The mathlib Community, 2020). These courses "require students to master three different languages: informal mathematical language, formal symbolic logic, and a computational proof language that lies somewhere in between" (p. 1). Instructors aim to focus students' attention on "appreciating the relationships between them and the multiple representations [they each] support" (*ibid).* Preliminary research into the efficacy of such a course demonstrated that students can succeed in these aims and that, while the appreciation for mathematical proof generated by these courses differs from that generated by more traditional instruction, students frequently end the course with a more holistic and sophisticated view of the field of mathematics.

At present, ATPs are used as teaching aids, rather than a singular mode of instruction. However, with recent innovations in computer-aided assessment and fully autonomous course structures for calculus and algebra, a fully autonomous introduction to proof MOOC looks increasingly possible.

Gratwick et al. (2020) presented a fully autonomous course on undergraduate algebra and calculus, "interleaving textbook-style exposition with videos of worked examples, interactive applets, and practice questions". Students' responses to practice questions and assessments were graded using STACK – a computer algebra system capable of providing personalized feedback based on the mathematical properties of student-produced responses to practice questions and examinations (Sangwin, 2013). While this software is not yet capable of assessing students' understanding of mathematical proof, researchers are beginning to take steps in this direction, building on existing assessment protocols to develop fully autonomous equivalents based on, for example, faded worked examples, automated checking of student-produced (counter-)examples, and multiple-choice reading comprehension tasks (Bickerton & Sangwin, 2020). Coupled with the increasing pedagogical potential of Automated Theorem Provers, and the increasing sophistication of artificial intelligence used to translate between formal and natural languages

(Brunello et al., 2019), the valid automation of proof comprehension assessment seems to be a question of *when* not *if*.

Opportunities
The increased scope of fully autonomous instruction in mathematics, paired with the increasing global access to computing and the internet, produces a new potential for independent study. This novel pathway toward a degree in mathematics holds the potential to improve access to tertiary mathematics education for vast swaths of the population previously unable to fund a traditional degree in tertiary mathematics.

Moreover, ATPs have the potential to provide insights into the mathematical world that simply are not available 'by hand' (Avigad, 2019). Both instructors and students reported on this potential value to mathematics education, leading a new generation of mathematicians whose understanding of the values and norms of the mathematical community are inextricably linked to their understanding of formal logic, computer programming and the ability to translate between formal and informal representations of mathematical objects.

Dangers to community, equity, and epistemic injustice
While the benefits of automated instruction are numerous, we must not lose sight of the dangerous consequences of such change. First, mathematics is an inherently social pursuit and as such, learners of mathematics are often most successful when surrounded by other learners. With the potential of automating the entire tertiary experience in mathematics, comes the potential of denying students' the community of co-learners we know to be so pivotal to their apprenticeship in the discipline. Second, and perhaps more importantly, are questions of equity and access. Equity in online instruction is often discussed through frames of access and inclusion (Gutiérrez, 2009), but we must also attend to the potential injustices perpetuated by automation. The dangers of unfettered automation have been studied in, amongst others, the insurance algorithms, driverless cars, and employment opportunities (Noble, 2018). Similar topics related to accountability and unconscious bias are likely raised by the automation of instruction and assessment in mathematics.

Final remarks
This essay presents the potential of existing ideas in the field of upper tertiary mathematics, and highlight its dangers if introduced as the product of societal inertia, rather than careful and deliberate planning. Recent technological advances represent a limitless potential for the democratization and distribution of tertiary education. But in this increasingly algorithmic world, we must carefully and deliberately make space for cultures, practices, and ways of being that might easily be swept away.

References.
Avigad, J. (2019). Learning Logic and Proof with an Interactive Theorem Prover. In G. Hanna, D. Reid, & M. de Villiers (Eds.), *Proof Technology in Mathematics Research and Teaching: Mathematics Education in the Digital Era* (pp. 277–290). Springer, Cham.

Bickerton, R. T., & Sangwin, C. J. (2020). *Practical Online Assessment of Mathematical Proof*. https://arxiv.org/abs/2006.01581

Brunello, A., Montanari, A., & Reynolds, M. (2019). Synthesis of LTL formulas from natural language texts: State of the art and research directions. *Leibniz International Proceedings in Informatics, LIPIcs*, *147*(17), 171–1719.

Dawkins, P. C., & Weber, K. (2017). Values and norms of proof for mathematicians and students. *Educational Studies in Mathematics*, *95*(2), 123–142.

Gratwick, R., Kinnear, G., & Wood, A. K. (2020). An online course promoting wider access to university mathematics. In Marks, R. (Ed.), *Proceedings of the 40th Conference of the British Society for Research into Learning Mathematics, 40(1).* https://bsrlm.org.uk/wp-content/uploads/2020/05/BSRLM-CP-40-1-04.pdf

Gutiérrez, R. (2009). Framing Equity: Helping students "play the game" and "change the game." *Teaching for Excellence and Equity in Mathematics*, *1*(1), 5–7.

Noble, S. (2018). *Algorithms of oppression*. New York University Press.

Sangwin, C. J. (2013). *Computer-aided assessment of mathematics*. Oxford University Press.

The mathlib Community. (2020). The lean mathematical library. *CPP 2020 - Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, Co-Located with POPL 2020*, *Cic*, 367–381.