



© 1997–2009, Millennium Mathematics Project, University of Cambridge.

Permission is granted to print and copy this page on paper for non-commercial use. For other uses, including electronic redistribution, please contact us.

25/11/2008

News

Automated mathematics



Does $1 + 1$ really equal 2 ? Only mathematicians could ask such a strange question and only mathematicians could (and, as you'll be relieved to hear, have) come up with a rigorous proof based on a meticulously worked-out definition of the whole numbers. Given this insistence on formality, it may be surprising that your average proof in an academic journal is riddled with informalities. Authors gloss over details, appeal to pictures, even intuition, and take hidden leaps of logical faith that, philosophically speaking, aren't entirely justified. These days mathematics contains proofs so long and complex that few people are able to check and understand them in full, yet once a result has made it through the peer review process and into a journal, its truth is taken as read.

Soon a task for computers?

All this is a far cry from the mathematical dream which started with Euclid over 2000 years ago: that every mathematical statement should be derived from the very axioms of mathematics in a sequence of verifiable logical steps. Proofs which do this are known as *formal proofs*, and they are the focus of a special issue of the *Notices of the American Mathematical Society*, which is now [freely available online](#). The issue contains four accessible articles exploring the current state of the art of formal proofs and giving some practical guidance for those feeling the need to formalise their maths.

Automated mathematics

The reason why you'll find few formal proofs in the mathematical literature is that they are tedious and long. Famously, it took Bertrand Russell and Alfred Whitehead until page 86 of the second volume of their *Principia Mathematica* to prove that 1 plus 1 is indeed 2, a result they accompanied by the comment "the above proposition is occasionally useful". The definition of the number 1, when laid out according to the strict rules of the *Bourbaki* school of mathematics, would run to over four trillion symbols. Even if mathematicians submitted to the tedious task of expanding all their proofs in such a painstaking manner, it would not be humanly possible to check all the logical steps involved. But what mathematicians are currently exploring is whether the essentially mechanical task of checking formal proofs could be left to those who are best at the mind-numbing application of rules: computers.

Computer assisted proofs have been a hot topic in mathematics for a while now. Debate started in the 1970s when the mathematicians Kenneth Appel and Wolfgang Haken produced a proof for the *four colour theorem* which relied on a volume of computer calculations so large that no human could possibly check them all. Given the amount of bugs that routinely creep into computer programs (according to one of the articles in the *Notices*, a typical software program contains one bug per 100 lines of code) this raises the question whether such computer assisted proofs can be trusted, and their validity is indeed hotly disputed by mathematicians.

The computer tools that have been developed to help with formalising proofs, the aptly named *computer proof assistants*, are essentially different from purpose-built programs like those used by Appel and Haken. Rather than performing a set of specific calculations, the proof assistants can take as input any formalised mathematical proof and check its validity. Several proof assistants are currently competing for mathematicians' favour. Each is based on the basic axioms of mathematics (what exactly these axioms should be is disputed, but there is a general consensus) and rules of logical inference. Once a mathematical proof has been formalised and translated into a language the proof assistant can understand, the assistant goes off to check that there are no holes, and that the result does indeed derive directly from the axioms. The hope is that with the help of these tools, many important mathematical results can eventually be made absolutely water tight.

The question of bugs does of course remain. How do you know that the assistant doesn't contain any? Thomas C Hales, a proponent of computer proof assistants writing in the *Notices*, points to the surprising simplicity of the central bits of code that drive the assistants: one of them, known as *HOL Light*, is implemented by a section of code a mere 500 lines long. The code is openly available for scrutiny and has passed under the eyes of expert logicians. What's more, proofs can be verified by more than one assistant, and the chance that several assistants erroneously verify the same false proof is extremely low. In fact, Hales believes that "the error rates of top-tier theorem-proving systems are orders of magnitude smaller than error rates in the most prestigious mathematical journals."

Currently, formalising a proof and having it checked by a computer proof assistant is still an enormous undertaking, and most working mathematicians have little to do with computerised theorem proving tools. However, the idea is that, as the library of formally verified proofs grows, mathematicians will routinely interact with computerised systems, submitting their own proofs and checking which results have already been verified. Hales hopes that the system will eventually "become a familiar part of the mathematical workplace, much as email, ... computer algebra systems and web browsers are today."

So much for checking the work of human mathematicians, but can computers actually come up with their own proofs? Some progress has been made in this area, too, especially with programs designed to solve special classes of problems. General theorem discovery tools, however, operate at most at the level of your average undergraduate exercise, proving, for example, that the number e is irrational. A sizeable task for an undergraduate, but disappointing viewed from the cutting edge of mathematical research. For the moment, and probably for some time to come, mathematicians' jobs are safe from their automated counterparts.

Further reading

- The articles in the *Notices of the American Mathematical Society* are freely available online;
- You can read about the four colour theorem in the *Plus* article The origins of proof;
- And you can find out more about the controversy surrounding computer assisted proofs in Welcome to the maths lab.

[Post a comment on this story](#)

Marianne Freiberger



Plus is part of the family of activities in the Millennium Mathematics Project, which also includes the NRICH and MOTIVATE sites.