Record: 1

Title:	HOW TO (REALLY) TRUST A MATHEMATICAL PROOF.
Authors:	Rehmeyer, Julie
Source:	Math Trek; Nov2008, p2-2, 1p
Document Type:	Article
Subject Terms:	MATHEMATICS COMPUTER science MATHEMATICAL proofs PROOF theory GONTHIER, Georges MICROSOFT Corp.
Geographic Terms:	CAMBRIDGE (England) ENGLANDReport Available
Abstract:	The article discusses mathematical develop computer proof-checking systems. A group of mathematicians and computer scientists believe that with new proof-validation programs, the dream of a fully spelled-out, rigorous mathematics, with every deduction explicit and correct, can be realized. In 2004, Georges Gonthier, a computer scientist at Microsoft Research in Cambridge, England, verified the proof of the four-color theorem by computer.
Lexile:	1210
Full Text Word Count:	1160
Accession Number:	35833288
Database:	Middle Search Plus

HOW TO (REALLY) TRUST A MATHEMATICAL PROOF

Mathematical develop computer proof-checking systems in order to realize century-old dreams of fully precise, accurate mathematics.

The one source of truth is mathematics. Every statement is a pure logical deduction from foundational axioms, resulting in absolute certainty. Since Andrew Wiles proved Fermat's Last Theorem, you'd be safe betting your life on it.

Well … in theory. The reality, though, is that mathematicians make mistakes. And as mathematics has advanced, some proofs have gotten immensely long and complex, often drawing on expertise from far-flung areas of math. Errors can easily creep in. Furthermore, some proofs now rely on computer code, and it's hard to be certain that no bug lurks within, messing up the result.

Bet your life on Wiles' proof of Fermat? Many mathematicians might decline.

Still, the notion that mathematical statements can be deduced from axioms isn't hooey. It's just that mathematicians don't spell out every little step. There's a reason for that: When Bertrand Russell and Alfred North Whitehead tried to do so for just the most elementary parts of mathematics, they produced a 2,500-page tome. The result was so difficult to understand that Russell admitted to a friend, "I imagine no human being will ever read through it."

Where humans falter, computers can sometimes prevail. A group of mathematicians and computer scientists believe that with new proof-validation programs, the dream of a fully spelled-out, rigorous mathematics, with every deduction explicit and correct, can be realized.

Indeed, Freek Wiedijk of Radboud University Nijmegen in the Netherlands says a revolution is already occurring. He writes in the December Notices of the American Mathematical Society that in the future, "most mathematicians will not consider mathematics to be definitive unless it has been fully formalized."

The first proof-validation programs were created more than 20 years ago. Until recently, though, they were so cumbersome that the only users were the researchers who had created and were trying to improve them. Furthermore, even those researchers were only tackling relatively simple theorems. In the last five years, though, those users have finally been able to verify some remarkably complex and

difficult proofs. Before long, they say, ordinary mathematicians will be using these tools as part of their everyday work.

Perhaps the most remarkable success so far came in 2004, when Georges Gonthier, a computer scientist at Microsoft Research in Cambridge, England, verified the proof of the four-color theorem by computer. The problem dates back to 1852, when a college student noticed that only four colors were needed to fill in a map of the counties in England such that no adjacent counties shared a color. It took until 1976 to mathematically prove that four colors were enough for any map. That proof was more than 500 pages long and relied on computers to check nearly 2,000 special cases. Many mathematicians objected to the proof because it was impossible to check by hand.

Gonthier used a proof-checking software package to formalize the entire proof, reducing both the text and the software for the special cases to an enormously long series of simple deductions.

Of course, if the proof-checking software itself has bugs, Gonthier's verification of the four-color theorem itself could be invalid. To guard against this possibility, the designers of the proof-checking software make the "kernel" of code that implements the axioms and rules of inference as short and simple as possible. One program, HOL Light, has fewer than 500 lines of code in its kernel, few enough that humans can check it by hand. John Harrison, the creator of HOL Light, has also checked the code using other formal proof-checkers!

The software may not be able to produce perfect certainty, but Thomas Hales, a mathematician at the University of Pittsburgh, calls the four-color theorem "one of the most meticulously verified proofs in history."

Hales is one of the first working mathematicians to embrace the proof-checkers, because he ran up against the limits of mathematical certainty himself. He proved the Kepler Conjecture in 1998, which is another theorem that is simple to state but remarkably hard to prove. The Kepler Conjecture says that the pattern grocers use to stack oranges packs the most oranges into the smallest space. As with the four-color theorem, Hales used a computer to check many, many special cases, and the proof consisted of 300 pages of text and 40,000 lines of computer code.

When he submitted his result for publication, he received only a qualified acceptance. The letter from the editor explained that "the referees put a level of energy into this that is, in my experience, unprecedented." Nevertheless, the referees ended up only 99 percent certain that the proof was correct. The referees were unable to check the computer code at all.

Hales decided that 99 percent certainty wasn't good enough for him. He started the "Flyspeck" project (named from the acronym FPK, for Formal Proof of the Kepler Conjecture) to formalize his entire proof. When he began, he estimated that it would take 20 person-years to complete it (i.e., one person working for 20 years, for example, or 10 people working for two years). He says now that he is about halfway through, and his team has indeed devoted about 10 person-years.

Hales and Gonthier are managing to do more than simply check those particular proofs. In the process, they are creating a library of basic formalized results other mathematicians can use to formalize new proofs. Since new proofs always rely on many, many previous proofs, this library provides the essential foundation mathematicians need to efficiently use the proof-verification software programs in their daily work.

Once that library is created, widespread use may not be so far off. Gonthier has been surprised to find that with experience, coding a proof takes little more effort than typesetting an ordinary mathematics article, which mathematicians do regularly. "The actual coding of results seems to go on pretty quickly," Gonthier says.

The hard part is the early stages, he says, teaching the computer what an early graduate student would know. Mathematicians use many, many tiny results and methods that they never write down explicitly. "Most of what you learn from a textbook is in the exercises," he says. "An entire part of the theory is something never described literally. If you want to formalize a theory, you have to find a good description for these things."

Gonthier believes that ordinary mathematicians may start formally verifying their proofs within the decade. Cameron Freer of the Massachusetts Institute of Technology is beginning a collaborative

project called Vdash that he hopes will inspire many mathematicians to pitch in and help build the basic library of results. Hales warns, though, that this will be a formidable task. "To undertake the formalization of just 100,000 pages of core mathematics would be one of the most ambitious collaborative projects ever undertaken in pure mathematics, the sequencing of a mathematical genome," he writes.

~~~~~~~~

By Julie Rehmeyer

~~~~~~~~