Abstract of the 15th Annual Gödel Lecture

▶ MICHAEL O. RABIN, *Proofs persuasions and randomness in mathematics*.
Division of Engineering and Applied Sciences, Harvard University, Pierce Hall, 29 Oxford St., Cambridge, MA 02138, USA.
E-mail: rabin@deas.harvard.edu.

It is customary to hold up mathematical proofs as ideal models of certainty. In principle a proof can be completely formalized and annotated and then it can be handed to other mathematicians and logicians who can "automatically" and efficiently verify it beyond any doubt. In practice the establishment of a mathematical statement is a social process. Someone proposes a proof, he and others check it and errors are often uncovered and corrected, until finally the community of mathematicians accepts the statement and its proof as correct.

Over the past three decades computer scientists have injected a number of revolutionary ideas and methods into the realm of proofs. These innovations mainly involve the use of randomness, an idea that flies in the face of the proverbial absolute certainty of mathematical proofs. Another salient feature of these novel methods of proofs is that in a well-defined sense they constitute self-persuasions rather than proofs. We shall discuss three outstanding examples of these innovations.

Randomized or probabilistic proofs. Consider the statement: $n = 2^{400} - 593$ is a prime number. You prove it by randomly choosing 100 integers $a_1, \ldots, a_{100}$, and performing on each an easily computable test $W_n(a_j)$. If all these tests produce the truth value F, you declare n to be prime. If n is composite rather then prime, the probability of erroneously concluding that it is a prime is smaller than $(1/4)^{100}$. Now, first we have established a mathematical fact with a probability of error, albeit provably a very small one. Second, the proof is non-transferable. If you hand the integers $a_1, \ldots, a_{100}$, to somebody else and he verifies that $W_n(a_j) = F$ for $1 \leq j \leq 100$, it does not prove anything for him. Namely, n is perhaps in reality composite and you have deliberately chosen the integers a_j in order to mislead him. A person must persuade himself of the primality of n by the randomized method, by effecting the random choice of the a_js on his own. This raises additional profound questions as to what randomness means in the real world and whether, assuming that we know the meaning of randomness, random processes exist in nature.

Zero Knowledge and Interactive Proofs. Assume that P know, for a propositional formula $A(p_1, \ldots, p_k)$, a satisfying truth value assignment for the propositional variables $p_1, \ldots, p_k$. The Prover P can cooperate with a Verifier V in an interactive Zero Knowledge Proof (ZKP) that he, the Prover, knows a satisfying truth value assignment for the formula $A(p_1, \ldots, p_k)$. The Verifier V poses to P a small number m of randomly chosen challenges. If P correctly responds to all challenges, then V is persuaded that P knows a satisfying truth value assignment. The probability that V was mislead is smaller the $(1/2)^m$. The Zero Knowledge aspect of the proof means that besides being persuaded of the Prover's claim of knowledge, V learns *nothing*, not just about the assignment but about anything else. Hence the name Zero Knowledge Proofs. The theory of ZKPs provides precise and convincing definitions for the concepts of Zero Knowledge and of proof of knowledge by P, in addition to the surprising fact that ZKPs are possible.

The method of interactive proofs has other surprising applications. A computationally powerful (in a precisely defined sense) Prover can provide to any Verifier an interactive proof that a given quantified prepositional formula is true. This has far reaching, albeit purely theoretical, consequences.

Probabilistically Checkable Proofs (PCPs). Assume that P has, as before, a satisfying truth value assignment for the propositional variables $p_1, \ldots, p_k$, of a prepositional formula $A(p_1, \ldots, p_k)$. He can effectively, i.e., by a polynomially long computation, transform the

formula A into a formula $B(q_1, \ldots, q_m)$ so that A is satisfiable if and only if B is satisfiable. P then writes satisfying truth values $v_1, \ldots, v_m$ into, say, a computer memory. Now, the verifier V can randomly choose a small number, say 20, memory locations, read the stored truth values (i.e., either F or T) and perform a very simple test. If the test comes out correctly, V concludes that B, and hence A, is satisfiable. The probability that he will be mislead is smaller than $(3/4)^{20}$. The startling aspect is that the number of truth values or bits that he reads for this verification is *independent* of the size or number of variables of the formula A. By some transformations we can conclude that one can convince himself of the existence of a proof of Fermat's Last Theorem by examining just 20 randomly chosen bits of a certain formalized form of the proof written into memory by A. Wiles.

Computer generated proofs. In the past two decades proofs of some significant results were generated by computers, or at least by man-computer cooperation, where the computer has played a pivotal role. Examples are the proof, due to Appel and Haken, of the four-color conjecture by computer, and the extensive work by Doron Zilberger. We shall discuss the meaning and implications of these developments.