Mathematicians Reach Factoring Milestone

The international fraternity of mathematicians chips in to achieve a notable first: factoring the first 100-digit number

CHALK ANOTHER ONE UP for the electrons: $11^{104} + 1$, the first "hard" 100-digit number, has been fully factored. After 3½ weeks of number crunching, Mark Manasse at the Systems Research Center of the Digital Equipment Corporation (DEC) in Palo Alto and Arjen Lenstra of the University of Chicago set the new record in factoring, on 11 October, at 2 o'clock in the morning (California time)—an appropriate hour, since nearly all of the computation was done in the middle of the night.

More significant than the 100-digit milestone is the fact that the work was carried out by a worldwide network consisting of hundreds of computers in the United States, Europe, and Australia, each working part time on a piece of the problem, and communicating by the existing network of electronic mail. The coordinated effort has implications for the ostensible security of the "unbreakable" public-key cryptosystems based on the difficulty of factoring large numbers. "This shows that it's not so difficult to mobilize resources," says Andrew Odlyzko of Bell Labs in Murray Hill, New Jersey, who contributed time on three machines.

Manasse notes that the work was done during otherwise idle time on the computers, and thus was done at essentially no cost. At DEC the work was distributed using a program written by John Ellis, also of the Systems Research Center, that hunts down idle workstations. Manasse believes that a low priority, loosely coupled network approach can be applied not just to factoring, but to problems in computational chemistry, meteorology, and linear programming as well. "There's a whole bunch of computers out there that don't do much during the night," he says.

Factoring a number consists of finding the prime numbers that divide it. For instance, 105 factors into $3 \times 5 \times 7$. There is an utterly straightforward process for factoring any number N, namely trial and error: simply try dividing N by 2, 3, 5, 7, and so forth. (You can stop if you reach $\sqrt{N}$, since N cannot be the product of two numbers both larger than $\sqrt{N}$.) However, this proceure is grossly inefficient if N has large prime factors (or turns out itself to be prime): a billion computers each doing a billion trial divisions per second would still take roughly 10^{23} years to factor the product of two 50-digit primes.

Oddly enough, it is extremely easy to prove that a number is factorizable without actually factoring it. Every prime number p enjoys the property of dividing anything of the form $a^p - a$. For instance, 5 divides $2^5 - 2 = 30$. Thus if, say, $2^N - 2$ is not divisible by N, then you know that N is factorizable, but you do not know any of its factors. There are efficient ways to implement this test for primality.

As it turns out, an occasional nonprime will pass this test—mathematicians call such numbers "pseudoprimes"—but other more sophisticated tests can expose the impostors. The upshot is that primality testing is relatively easy, whereas factoring still appears to be very difficult.

"There's a whole bunch of computers out there that don't do much during the night."

Numerous factoring methods that improve on trial and error have been introduced over the years, with the most dramatic advances coming in the last 10 years. The method used to crack $11^{104} + 1$—which has "small" factors of 2, 17, and 6,304,673 (whose product equals $11^8 + 1$)—was invented by Carl Pomerance of the University of Georgia at Athens, and is known as the "quadratic sieve."

The idea behind the sieve is to find numbers X and Y such that $X^2 - Y^2$ is divisible by N. If neither $X - Y$ nor $X + Y$ is divisible by N, then the greatest common divisor of $X - Y$ (or $X + Y$) and N—easily computed by the Euclidian algorithm—is a factor of N.

The sieve works by finding a collection of numbers x for which the remainder of x^2 under division by N is easily factored into small primes belonging to a preestablished "factor base" for N. (The factor base merely consists of small primes, not the primes that divide N. For $11^{104} + 1$ the factor base had 50,000 primes in it.) It does this by searching through a larger collection of numbers x and "sieving out" those with the desired property. This step, which constitutes the bulk of the computation, lends itself to distributed processing—many computers can be put to work on different ranges of values of x. Moreover, their efforts need not be synchronized: a "master" computer simply waits until all the reports are in. "People came in when they felt like running the program," Manasse says.

Once the sieving is done, the master computer does a "simple" matrix calculation to determine a subset of the collection of numbers x such that the product of the corresponding remainders—whose factorization is completely known—has all even exponents in its prime factorization, hence has the form Y^2. (A matrix is simply a rectangular array of numbers. For the quadratic sieve, each horizontal row corresponds to a number x and each vertical column corresponds to a prime p in the factor base; the numerical entry in "row x" and "column p" is simply the exponent of p appearing in the remainder of x divided by N. More precisely, the entry is 0 if the exponent is even, and 1 if the exponent is odd. The matrix calculation amounts to finding a set of rows that when added together results in all even numbers for the exponents. Keep in mind, the matrix for $11^{104} + 1$ had 50,000 columns and at least that many rows.)

With X denoting the product of the corresponding x's, it is an elementary fact that $X^2 - Y^2$ is divisible by N. The process stumbles only if $X - Y$ or $X + Y$ is divisible by N, in which case the matrix is returned to for another subset of x's, or a few more x's are sieved. Heuristically, there is at least a 50–50 chance that the process will work for any subset of x's. In the case of $11^{104} + 1$, Manasse says, the first attempted subset was "unlucky," the second produced the factorization, the next few were again unlucky, and then another produced the same factorization, indicating that the factors—a 41-digit number and a 60-digit number—were primes. The network is now tackling a 102-digit number and has its eye on a number of 106 digits.

The difficulty of factoring large numbers, especially those formed by multiplying two large prime numbers together, is at the heart of a public-key cryptosystem that was proposed in 1977 by Ronald Rivest, Adi Shamir, and Len Adleman at Massachusetts Institute of Technology. Factorization is not known, however, to be intrinsically dif-

ficult, and as factorization algorithms improve, such cryptosystems are placed in jeopardy.

"If you had a sufficiently dedicated adversary, then even numbers as large as 512 bits (154 digits) probably will not be safe for more than another 3 to 4 years," Manasse says. In the absence of radically new factorization methods, though, cryptographers can respond simply by keeping ahead in the digit race. The only real danger is to secrets that must remain secret for more than a few years.

In 1977, Rivest and his co-workers published an "unbreakable" cipher based on a 129-digit number, with a $100 prize offered for its solution. At the time, factoring a 40-digit number was an accomplishment. But 129-digit numbers are only about 100 times harder to factor than 100-digit numbers. Manasse expects to be able to upgrade the current network easily by a factor of 10, and possibly by a factor of 100.

Asked if he thinks he can win the $100, Manasse says, "If I do I'll have to split it a whole bunch of ways." ■ BARRY A. CIPRA

Barry A. Cipra is a mathematician and writer based in Northfield, Minnesota.

Discovery Gets a Clean Bill of Health

NASA has made hundreds of fixes to the shuttle; the evidence from Discovery is that virtually everything worked

WHENEVER ENGINEERS have to go into an already complex piece of equipment and make hundreds of modifications, as the National Aeronautics and Space Administration (NASA) did to the space shuttle system in the aftermath of the 1986 Challenger accident, a wise person has to be concerned: there is always a chance that the fixes will create more problems than they solve.

Thus the institutional sigh of relief at the dramatically successful flight of the space shuttle Discovery on 29 September to 3 October, a flight that marked the nation's return to manned space flight after a 32-month hiatus. Not only was the mission itself nearly flawless, but all the post-flight indications are that every piece of modified equipment performed as expected.

"With all the changes it's incredible how few problems we've had," says Joseph E. Mechelay, manager of the Flight Data and Evaluation Office at NASA's Johnson Space Center near Houston. "The results of this flight indicate that we haven't screwed anything up."

Indeed, NASA treated this mission, the 26th launch of the shuttle since flights began in 1981, as if it were the test flight of a brand new vehicle. And now that Discovery is back on the ground, it is being examined in minute detail. Some highlights:

■ **The solid rocket boosters.** These are the huge white crayons on either side of the shuttle's rust-colored external fuel tank. Rather like skyrockets in a fireworks display, they fire for the first 2 minutes after liftoff and then fall away into the ocean, where they are retrieved for reuse.

In the last, fatal flight of Challenger on 28 January 1986, however, a design flaw allowed exhaust flame to burn through a rubber O-ring gasket in one of the boosters, and then to escape from the side through a joint in its metal casing. As a result the booster tore loose from its mount 73 seconds after launch and destroyed the whole vehicle. Much of the 2½ years since then has been spent in testing and validating a highly modified design for the boosters. In fact, says one manager at Morton Thiokol, the Utah-based company that builds the boosters, about four to five times as much work went into this effort as went into the original development in the mid-1970s.

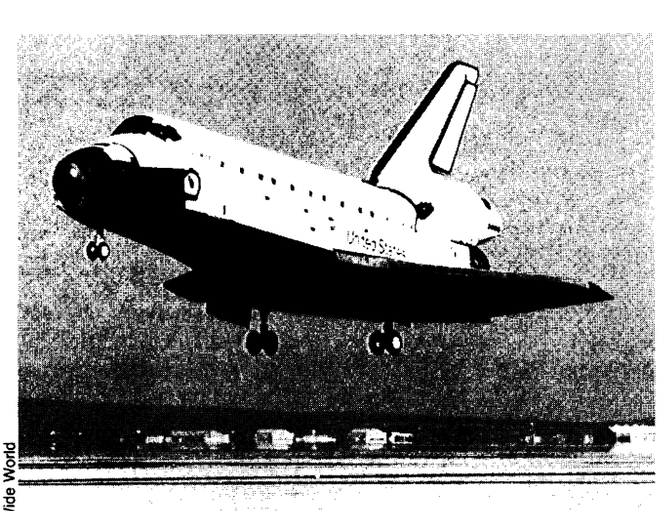
Discovery's boosters are now undergoing a preliminary examination at Florida's Kennedy Space Flight Center. "All indications are that they worked as planned," says Russell Bardos, head of NASA's shuttle propulsion office. "A cursory look shows no anomaly at all."

"It doesn't appear that any gas got to any O-rings," agrees Myron Uman, executive director of an ad hoc committee set up by the National Research Council to provide independent oversight of the booster redesign.

Once the preliminary inspection is completed in Florida the boosters will be sent back to Morton Thiokol in Utah, where they will be torn down, examined in detail, and then refueled for another flight.

■ **The main engines.** These are the three large cones located in the tail of the orbiter. During launch they burn some 780 tons of liquid hydrogen and liquid oxygen fuel located in the external tank. They produce more thrust per kilogram of weight than any engines ever built. They have been plagued with problems from the beginning, notably with bearing wear and cracks in the engines' many weld joints. During the hiatus, NASA therefore made some 35 upgrades to the engines. The agency also tightened up on inspection and certification procedures, and inaugurated the most aggressive ground testing program in the history of the main engines. Engines under test are routinely fired for more than 2000 seconds at a stretch, for example, about five times as long as they will be fired during an actual flight.

The data from Discovery suggest that the rigor paid off. Telemetry showed no indication of any problem with the engines during ascent, and a preliminary inspection on the ground shows none of the problems that



Wide World

Touchdown. *After 4 days in orbit, the space shuttle Discovery makes its final approach for a landing at Edwards Air Force Base in California.*