The endless search for primality

Is there more to the continuing search for prime numbers than the curiosity that sustains the enthusiasm of stamp-collectors? The discovery of the largest so far (at Harwell in Britain and by Cray Research) refers.

THE discovery of the largest prime number so far, herewith made public, may not set the world alight, but is well worth modest celebration, if only because of the circumstances of its discovery. Briefly, it appears that Cray Research provides its customers with a program for telling whether numbers are true primes. Naturally, a purchaser who has just written a cheque for several million dollars will be comforted if he can check that his new purchase verifies the primality of a known prime number, perhaps one chosen at random from a list of known primes.

In this case, the customer is the Harwell Laboratory of what is now called AEA Technology, previously the principal research laboratory of the UK Atomic Energy Authority. The computer is a Cray-2 machine equipped with four central processors. Like others before them, the Harwell people set out to test the primality of large Mersenne numbers, which have the form $2^q - 1$, where q is itself a prime number. What they have found is that the Mersenne number with $q = 756{,}839$ is also prime. It is much larger than that which previously held the record, with $q = 216{,}091$.

The sheer size of these numbers is awesome. In decimal notation, Harwell's new prime consists of 227,832 digits, almost exactly three times as many as the previous largest prime. A little mental arithmetic will show that it will never be published as ink on paper: if crammed with text, this page would hold 8,000 characters, so that Harwell's prime would occupy close on 30 pages of *Nature*. And, given that proof-reading would be so nearly impossible as to be pointless, only machine conversion from electronic signals into print would be practicable. (Similar problems arise with the nucleotide sequences of DNA, of course.)

But the discovery should not be scorned on that account. For one thing, the search for prime numbers has an abiding interest, not least because of the continuing uncertainty about their placement among the integers. Euclid, some time ago, showed that the number of primes must be infinite by the simple remark that adding to or subtracting 1 from the product of all consecutive prime numbers up to a certain maximum will yield two further prime numbers, which also goes to show why prime numbers often occur in pairs. (The product of the first three primes — 2,3 and 5 — is 30, and both 29 and 31 are primes.) In other words, given any sequence of consecutive primes, it is always possible to construct two that are larger.

But the density of primes diminishes as integers become larger. Handwaving shows that. The larger a number, the larger the number of prime numbers less than or equal to its square root and thus the greater the chance that it will be divisible by one of them. More precisely, according to Poussin in 1899, the number of primes less than a number N is, to a first approximation, $N/\ln N$, where "ln" stands for the natural logarithm. Taken literally, this means that the density of primes declines only slowly with increasing size. For example, the number of primes per thousand consecutive integers in the region of 10^9 would be related to that in the region of 10^6 in the ratio of 6:9, or 2:3.

It is also important that this approximation is only the first term in a slowly converging series, whose accuracy has never been stringently tested. One of the uses of very large prime numbers is that they may, when accompanied by a knowledge of the neighbouring primes, make it possible to test the Poussin formula and other guesses by people such as Riemann at the density of primes. But the likelihood of that will hang on people's energy, and the spare time of machines such as the Cray-2.

Meanwhile, the use of Mersenne numbers in the search for ever-larger primes seems unavoidable, and that is a remarkable story in itself, best told in Volume 2 of Donald E. Knuth's *The Art of Computer Programming*. Marin Mersenne was a monk who first showed in 1644 that numbers of the form $2^q - 1$ are prime only if q is prime (for otherwise the number is divisible by a smaller Mersenne prime) and who conjectured — nobody knows how — that the first prime numbers of this form are those in which q is one of 2, 3, 5, 7, 13, 17, 19, 31, 67, 127 and 257.

Knuth, who is a terrier for these things, records that by the 1880s, it had been shown that $2^{67} - 1$ is not a prime and, soon afterwards, that $2^{61} - 1$ is a prime, which led Mersenne's supporters to plead that he had mistranscribed "61" and "67". But then, unhappily, it turned out (in 1911) that $2^{89} - 1$ is a prime not on Mersenne's list, and (in 1914) that $2^{107} - 1$ is another. Then, in 1922, it turned out that $2^{257} - 1$ is not a prime.

The usefulness of Mersenne numbers in the search for larger primes now rests on a different property of integers of the form $2^q - 1$, that by a suitable interpretation of the representations of numbers by strings of binary digits ("0" or "1"), division by those numbers can be reduced to shift and complementation operations. Cray Research's diagnostic prime program is based on what is called the Lucas-Lehmer test for the primality of Mersenne numbers, based on a rigorously proved algorithm.

The procedure sounds more complicated than it is. If q is the exponent of the Mersenne number to be tested, the trick is to compute a sequence of numbers L_n by the rule that $L_{n+1} = (L_n^2 - 2) \bmod (2^q - 1)$. The initial condition is that $L_0 = 4$ and the test is embodied in the assertion that the Mersenne number is prime if (and only if) L_{q-2} is zero. The notation "mod" is simply the remainder after division by the Mersenne number. It is quite fun to work out a few examples with pencil and paper; Knuth gives a formal proof.

The hard work, even for a Cray machine, is the repeated calculation of squares — in this case, 756,839 of them. Quite soon in the iteration, the numbers L will be comparable in size with the Mersenne number under test, so that the computer space required to store the square requires twice as many digits, getting on for half a million (in decimal notation).

Making this efficient has been the work of David Slowinski, the computer scientist at Cray Research who designed the program. He says that he has used all the tricks of his trade, and that it nevertheless took nineteen hours of time on Harwell's Cray-2 to obtain a proof that $2^{756{,}839} - 1$ is prime. But there is a bonus: this Mersenne number (like 31 others) can be used to generate what is called a perfect number by multiplying it by the factor 2^q. A perfect number (which cannot be a prime) is one that is equal to the sum of its prime factors.

It is understandable that the people at Harwell and Cray are delighted with this outcome. It is a triumph of what Knuth would call the programmer's art. But is it more than stamp-collecting? That depends. The obvious need now is for a means of calculating primes in the neighbourhood of giants such as Harwell's, and then for using these to test the conjectures about the density of primes. And while the occurrence of primes may be peculiar, it seems that nobody has yet proved that a systematic procedure for their prediction is impossible. That should keep the armchair practitioners of modular arithmetic happy for years.

John Maddox