New Code Is Broken

An Israeli mathematician found a way to break the trapdoor knapsack code—one of the public key cryptosystems

One of the first public key cryptosystems ever to be suggested has now been broken by Adi Shamir of the Weizmann Institute. Shamir, who has been trying to break the code for years, finally figured out a way on 20 April. Although variations on this popular code still seem to be secure, Shamir's attack is leading some cryptographers to ask whether these variations will be the next to fall.

The idea of public key cryptosystems was proposed in 1976 by Martin Hellman, an electrical engineer at Stanford University, and his two students Whitfield Diffie, now at BNR in Palo Alto, and Ralph Merkle, now at Elxsi International in Sunnyvale. These were to be a completely new sort of code in which knowledge of how to encode a message would not reveal how to decode it. Each user would have an encoding key and a decoding key. He would publish the encoding key but keep the decoding key secret. Anyone could then use the encoding key to send this user a message but only the intended recipient could decode it.

Shortly after suggesting that such codes were possible, Merkle and Hellman came up with a specific example, called the Merkle-Hellman scheme or the trapdoor knapsack. Shamir, Ronald Rivest of Massachusetts Institute of Technology, and Leonard Adleman of the University of Southern California also proposed a scheme and it was these two schemes that sparked scientists' interest in cryptography and that led to a debate between scientists and the government over the national security implications of open cryptography research and publications.

Shamir has broken the original version of the Merkle-Hellman code. This code is based on a very hard mathematical problem, the knapsack problem, whose solution can take thousands or even millions of years of computer time. Merkle and Hellman's idea was to make a code that could not be deciphered unless the code-breaker solved a knapsack problem.

What Merkle and Hellman did was to construct knapsack problems in such a way that they knew how to solve them. But anyone who was not privy to the construction process would not, presumably, know how to get the solutions.

To solve a knapsack problem, a person has to decide which of a large group of numbers were added together to give particular sums. (It's as though you were given a knapsack filled with packages. You know the weight of the knapsack and the individual weights of a large group of packages that could be in the sack. You want to use this information to decide which packages are in the sack.) In general, the only way to do this is to try all possibilities and it is trying all possibilities that makes these solutions so time-consuming. But some cases are easy.

For example, if the original set of numbers is what is called a superincreasing sequence, meaning that each number in the sequence is greater than the sum of all the numbers preceding it, then mathematicians can easily decide which numbers of the sequence were added together to give particular sums.

> "Adi's paper is the first foot in the door."

If you know you have a superincreasing sequence, then you know the largest number in your sequence that is less than the sum you want to decompose must be a member of the decomposition. The reason is that the sum of all the numbers prior to that largest number are less than the largest number and so less than the sum in question. So, knowing that the largest number less than the sum must be part of the group of numbers that add up to the sum, you subtract that largest number from the sum and start the same sort of analysis over again. This is called the "greedy" algorithm.

Merkle and Hellman decided to use this special property of superincreasing sequences to make a code. A person would mathematically scramble a superincreasing sequence and turn it into a sequence that no longer appeared to be super-increasing. He would keep the method of scrambling (and unscrambling) secret. Then he would publish the scrambled superincreasing sequence and anyone who wanted to send him a message would use the published sequence to do so. Encoding a message entails adding certain numbers of the published sequence. Decoding entails decomposing those sums. Only the recipient of the message would know how to unscramble the published sequence and determine which numbers were added together to form the message.

The idea of using superincreasing sequences, says Ronald Graham of Bell Laboratories, "was a nice idea but it is the fatal flaw in their technique." The problem is that a large superincreasing sequence is an enormous spread of numbers—the smallest number in the sequence is much, much smaller than the largest one. Shamir was able to use the fact that the superincreasing sequence underlying the scrambled one cannot be completely disguised to convert the problem of breaking the code into a more approachable problem of integer programming. This step in itself is not so surprising, in retrospect, says Adleman. "Whenever mathematicians are confronted with a problem, their typical reaction is to manipulate it so that they can view it as a sort well studied in mathematics. Since there are arithmetic operations in this problem, it is natural to manipulate it into the well-studied form of integer programming."

Shamir's achievement was to realize how to solve the integer programming problem with a recently discovered fast method, called Lenstra's algorithm (*Science*, 3 April 1981, p. 31). Hellman remarks, "We looked at similar approaches and were not able to get them to work. But Adi knows more about integer programming than we do."

As soon as Andrew Odlyzko of Bell Laboratories learned of Shamir's work, he thought he knew a way to make the attack go even faster. Instead of using Lenstra's algorithm, Odlyzko reasoned, it should be possible to use continued fractions, which are used in number theory as a way of approximating real numbers with rationals. The advantage of using continued fractions, says Odlyzko, is that "there are well-known algorithms [for working with them] that can be implemented extremely rapidly. I've never heard of anyone programming Lenstra's algorithm although it is possible that it would be fast too."

Shamir says, "I was aware that Len-

0036-8075/82/0528-0971$01.00/0

Odd Amino Acids in a Meteorite

The controversy over whether any of the purely chemical processes that preceded the origin of life could have taken place in space is back again. Two geochemists, Michael Engel and Bartholomew Nagy of the University of Arizona, have reported* the discovery of the predominance of the L-form among amino acids in a piece of the Murchison meteorite. All amino acids in proteins have the L-structural form; amino acids synthesized in a test tube have equal proportions of the L-form and its mirror-image D-form.

Previous studies of Murchison and other organic-rich meteorites had reported only equal parts of D and L amino acids, termed racemic mixtures. Cosmochemists concluded from this that no selective processes, such as those that presumably chose the L-form for subsequent biological evolution, ever occurred in space. That conclusion must now be considered tentative.

Cosmochemists note the care lavished by Engel and Nagy on their analysis of the Murchison meteorite, but they are still uneasy about the results. They wonder why ratios of D- to L-forms of 0.2 to 0.7, as reported by Engel and Nagy, had not been seen before. Those ratios differ enough from 1 to have been detected by the methods applied to Murchison 12 years ago, according to Keith Kvenvolden of the U.S. Geological Survey in Menlo Park, California. He was a member of the group that first reported racemic mixtures of amino acids in Murchison.

Engel and Nagy offer several possible explanations for the conflicting results. One is that the particular stone that they analyzed, one of many collected from the same meteorite fall, had a different composition than the others. No one can rule out that possibility, but separate samples analyzed by several different laboratories have not shown any indication of such heterogeneity. Another possibility, Engel and Nagy say, is that the more severe extraction methods used in other laboratories could have converted some of the amino acids in the L-form to the D-form and produced a racemic mixture. That does not happen to amino acids added to samples in the laboratory, Kvenvolden notes, but he concedes that meteoritic organic matter may be in a physical or chemical form that could promote such racemization.

The primary concern of most researchers is terrestrial contamination. The Murchison meteorite fell on a sheep ranch in Australia, a site with obvious possibilities for organic contamination. But the usual signs of contamination are not there, Engel and Nagy point out. Amino acids that are common on Earth, such as tyrosine, methionine, and phenylalanine, are absent. They found only minor traces of serine and threonine, which are usually taken as signs of fingerprint contamination. In addition, the amino acids that are most tightly bound within the sample are also the least racemic, Engel notes, which would not be expected if the L-forms were a late, terrestrial addition.

What bothers some researchers the most is the curious behavior of some of the amino acids. The five nonracemic amino acids are all protein amino acids, some of the 20 that commonly form proteins. The two amino acids that seem to be racemic are both nonprotein—they are rarely or never found in proteins. One of these, α-amino-*n*-butyric acid (α-Aba), should behave as the five protein amino acids do, Kvenvolden notes, because all of their molecular structures permit transformations between L- and D-forms. If the excess L amino acids are not from contamination, the reasoning goes, there is no obvious way to generate extra L-form protein amino acids without affecting nonprotein amino acids such as α-Aba. "It's an enigma," says Edgar Hare of the Carnegie Institution of Washington, "there's no doubt about it." Engel suggests that one possible explanation is the generation of racemic α-Aba by the decomposition of some precursor.

Both Engel and Nagy emphasize that the possibility of contamination cannot be excluded, even though the sample meets traditional criteria for cleanliness. "Much more work is necessary to clarify the problem," says Nagy. He is not likely to get any arguments about that.—RICHARD A. KERR

Nature (London) **296**, 837 (29 Apr. 1982).

stra's algorithm may not be necessary but I included Lenstra's algorithm to show the solution is polynomial [meaning it can be implemented efficiently on a computer] even if you try to do a number of tricks. I tried to show that simple variants of the knapsack problem are not safe."

In the years since Merkle and Hellman proposed their scheme, they and others have suggested a number of modifications to make it more secure. Hellman says, for example, that he has always advised people to scramble the superincreasing sequence more than once. Graham and Shamir independently developed a variation on the scheme in which they add noise to the superincreasing sequence before scrambling it.

Shamir's attack does not as yet work on these variations of the Merkle-Hellman scheme. But, Shamir told *Science* in a telephone interview, "From now on it will be an infinite game." Each variation of the code will be subject to attack and, as each falls, a new variation will be proposed, Adleman agrees. "Adi's paper is the first foot in the door," he says. Odlyzko remarks, "Shamir's attack is very, very simple. Now that we see this simple solution I would not be surprised if other attacks are found as well."

Hellman says that on the basis of Shamir's result, he would advise anyone using the original Merkle-Hellman scheme to scramble the sequence more than once. Anyone who is contemplating using a variant of the code, Hellman suggests, should "wait at least a year and see what comes out. I certainly think it would be prudent to wait."

Interestingly, there have been persistent rumors that the National Security Agency has never considered codes based on the knapsack problem to be secure. AT&T was at one time considering using a knapsack code but, according to an informed source, the corporation consulted with the National Security Agency. On the advice of the agency AT&T reportedly chose to use a different sort of code.

As for now, it is too early to predict the implications of Shamir's result. Says Hellman, "At the very least, it's a clever piece of work."

Shamir is quite happy in part because he had hoped for some time to break the Merkle-Hellman code. In 1976, Merkle sent out fliers offering $100 to anyone who could break the code. Shamir kept a copy of the flier and mailed it to Merkle along with a seven-page abstract of his result. "Two days ago, I got my check for $100," Shamir told *Science*.

—GINA KOLATA

0036-8075/82/0528-0972$01.00/0