Automated Cryptography

Computers are a potent tool to codebreakers, but the ultimate decipherment still depends on the skill of clever humans

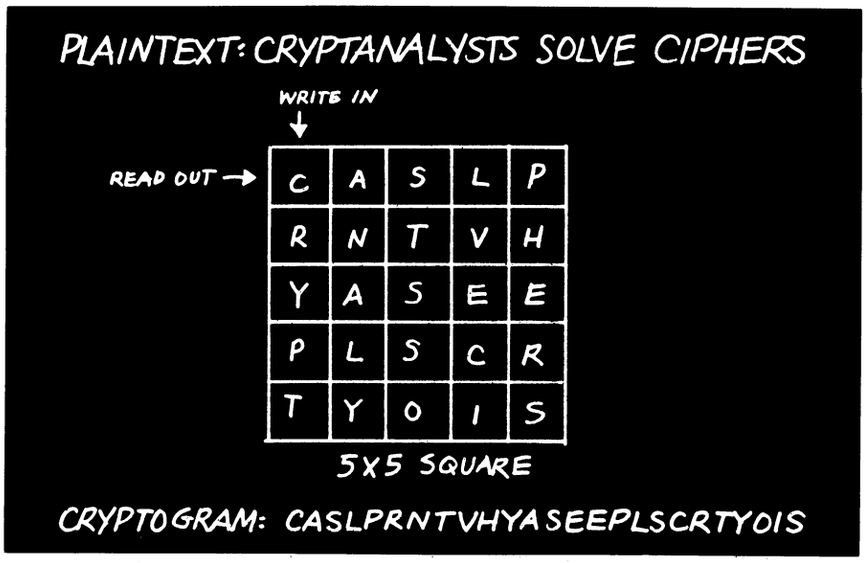
BY MICHAEL GUILLEN

Any child whose classroom note has been intercepted by a perturbed school teacher fully appreciates the security offered by even a simple code. In the sophisticated world of "international intrigue," the use of codes is motivated by similar worries. The major distinction, of course, is that today's national security installations utilize computers and mathematics to construct elaborate enciphering systems that are virtually impenetrable, simultaneously improving their chances of solving a rival code.

The superiority offered by computers has at no time been more publicly obvious than during the last several months. The National Bureau of Standards (NBS) is proposing the adoption of a certain mathematical prescription (algorithm) to generate a standard code. It would be commercially used for such purposes as protecting the privacy of personal and economic data currently flowing in and out of nationwide business computer complexes. Involved parties generally concede the proposed algorithm's ability to generate a code sufficiently complex to foil unethical tampering by rival interests.

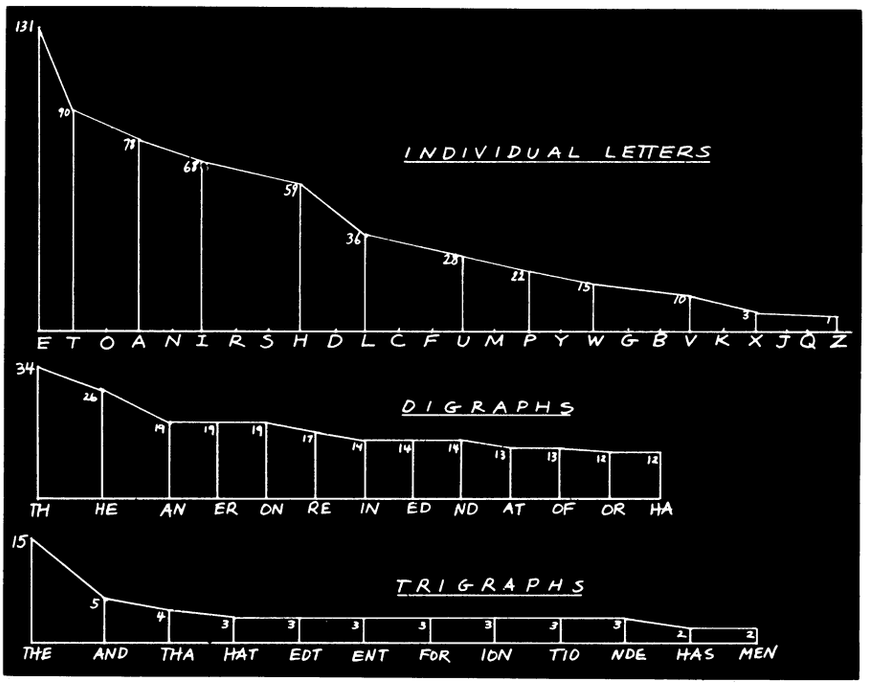
Many fear, however, that the mechanized superiority of the National Security Agency will allow it, alone, potential free access to the entire store of confidential information. Recent revelations by the House Select Intelligence Committee describe past NSA intrusions into private personal and business communications. Its official responsibility is to monitor and decipher international communications. The proposed "data encryption standard" will be the keynote subject of a cryptographic workshop at NBS scheduled for late September.

The role of the computer in contemporary cryptographic practices varies widely, depending on the application involved. Formally, however, the awe-inspiring computer capabilities are easily sabotaged by one of man's traditional weaknesses, "loose lip." Major security installations, like the NSA, therefore, cannot be persuaded to speak frankly about their business. Besides being a rather secretive sort, cryptanalysts are very fraternal, using a distinct language. The accompanying vocabulary list includes words that are freely used throughout the following text.



An example of how to generate a "columnar" transposition code.

An example of simple monoliteral substitution code used by Julius Caesar.



Normal frequency distribution of the letters of the alphabet (in uses per thousand).

Science Service, Inc. is collaborating with JSTOR to digitize, preserve, and extend access to *Science News*. www.jstor.org

Cryptanalysts' vocabulary

cipher: also cryptogram, an encoded message
digraph: any two-letter combination
frequency distribution (normal): a display of how relatively often in normal communication the letters of the alphabet are used; see diagram
general system: the overall technique employed to encode a message (see *key*)
invariant: a group of letters that always occur together (e.g. "QU" in English)
key: usually a single word that, together with a *general system,* must be known in order to decipher a message
plaintext: the uncoded message
probable word: a word that is guessed likely to be in a cipher based on the circumstances involved
polyalphabetic: usually with reference to a substitution code when more than one alphabet system is used to encipher a single plaintext
polygraph: any multiple-letter combination
substitution code: obtained when the plaintext letters are substituted by equivalents in another symbolic system
transposition code: obtained when the plaintext letters are retained but their order is systematically scrambled
word patterns: searched for in a cryptogram, a systematic repetition of a letter or group of them. A cryptanalyst, using special dictionaries (e.g. listing words according to length), may use these to help him solve a cipher

In certain ways, cryptography, with its impressive array of subtle tools to create and pick apart codes, remains a primitive art form. The computer has become a potent tool in the hands of the skilled cryptanalyst. Rather than replacing the human, the computer has freed him from the drudgery of repeated elementary operations often required to unravel a code. The ultimate decipherment, by and large, remains the responsibility of the cryptanalyst.

Most people have, at one time or another, toyed with the idea of inventing a coding system, the difficulty of which would stymie forever, even the most formidable of modern automata. The feat, as it happens, is not at all as difficult as some will boast it is. Most elaborate schemes made by amateurs with this purpose in mind are so impractical, however, even the intended recipient would spend half a lifetime deciphering a single message. Cryptograms, especially those encountered by agencies of national security, are often not directly soluble. This situation, in fact, is a hallmark of ultramodern cryptanalysis.

Before the advent of modern computers, the skill of one cryptanalyst was pitted against that of another. It was a "game" of wits, pure and simple. The purity of the activity is now often overrun by "combinatorial monstrosities." Very often, the solution of a code is retrieved only after testing a multitude of different letter combinations. Today, a cryptanalyst, armed with his computer, challenges the combined "wits" of another human and his machine. Wits for a computer, however, means brute force, the ability to perform millions of elementary operations in seconds. One analyst may program his computer to consume a billion steps to encipher a single *plaintext* message. He does so with the hope that an adversary's machine will lack the time and capacity to decipher it.

With computers, the task of code-solving grows in complexity at a breathtaking rate. By doubling the encoding capacity of a computer, you square the number of possibilities a would-be code-breaker needs to plow through in search of a solution. Doubling a computer's capacity can be as easy as adding a file-cabinet-sized memory bank to the existing complex. Furthermore, with the rapid evolution in the electronics industry, the attendant cost of such an embellishment is steadily decreasing. According to one estimate, computer costs are falling by 50 percent every five years.

In the midst of this mechanistic metastasis, national security installations often rely on "tricks" or "lucky breaks," either accidental or induced, to supplement the cryptographic efforts of their massive computer systems. Control Data Corp., known to manufacture custom-designed cryptographic machines for government purposes, markets a general consumer computer complex with trillions of bits (a unit of information) of storage capacity. The basic computer, known as the STAR model, can perform 100 million operations per second. This speed and capacity is a tremendous asset in solving many types of code.

Consider for a moment that you've been boastfully challenged by a friend to "break" one of his patented impervious creations. However ingenious the friend may be, his code will be one of two basic types: *substitution* or *transposition.* If cunning, he'll have used both in his message. In a substitution code, each letter of the alphabet can be assigned to one cryptic symbol (monoliteral) or several different ones (polyliteral). A transposition code is usually generated by first arranging the plaintext letters in a geometrical array (e.g. along columns) and then copying them back in altered order (along rows; see example). A transposition code can also be got by simply scrambling the order of the plaintext letters in any other systematic way (without using an array) that can be undone.

Returning to the challenge now, first list how often each symbol appears in your friend's cipher. Now, compare the shape of that *frequency distribution* with the *normal* one (see vocabulary list). If they match, there is a chance that the cipher might be just a simple monoalphabetic substitution (see example). By matching a cipher symbol with the letter that occurs with similar frequency, you will obtain a translation that should solve the code.

Assume the frequency distributions do not match up; your friend is a bit craftier. In this case, he may have used a polyliteral substitution code. This requires considerably more perseverance to discover the solution. The standard Morse code is polyliteral, a variable number of symbolic dots and dashes being equivalent to each letter. Sir Francis Bacon, the 17th-century natural scientist, used a polyliteral ciphering system based on five-letter groups of A's and B's. There being 32 different ways of arranging five A's and B's (2^5), there were more than enough combinations to accommodate the entire alphabet. (In Bacon's system, the word "man" might encipher as: AAAAB AABBA BBAAB, but without the spaces.) One way of solving this type of code is to try all the possible groupings of the cipher's letters. From among them, discover which arrangement yields the closest to a normal frequency distribution. That being found, the solution proceeds as before, using the resulting translation.

Merely scrambling the letters of a plaintext does not alter its normal frequency distribution. The problem in this case is to unscramble the letters, not translate them. To solve this type of transposition code, you need a lot of time. This is to test all of a cryptogram's possible "unscramblings," or permutations as they're called. From all the millions of possible letter combinations, only one will be the original message.

A complex code will successfully deceive not necessarily because it involves intellectual subtleties. It will thwart code-breaking efforts because the solution is suitably buried within millions of other, nonsensical, permutations. Therein, of course, lies the asset of today's computer. It can run through millions of permutations, search for revealing word patterns, try countless substitutions and critically compare combinations, all in a matter of seconds. With the aid of massive computers and elaborate programs, the human cryptanalyst has been relieved of tedious routine operations and given clues that significantly condense the list of possible solutions he needs to consider.

Depending on the application, human

LAUGH at the COLD!

It's 10°outside . . . Even getting colder. So you bundle up in layers and layers of heavy clothes. First with long underwear . . . then bulky, restrictive thermalwear on top.

Oh, you were warm, all right. Like in a Turkish bath. Because you began to perspire from all your activity. And perspiring in that mountain of clothes is like perspiring in a plastic bag! The perspiration is locked in. So there you are. Wet and miserable.

But now, at last, Damart has solved the problem. Because Damart invented underwear that keeps you *warm, dry and comfortable* no matter how cold it is or how long you stay out. Underwear that's soft and light so you can move easily. Underwear that *lets the perspiration evaporate through* so you always stay warm and dry next to your skin.

Damart does this with a new miracle fabric—Thermolactyl. It not only retains and reflects your natural body warmth, it's knitted to let *perspiration out!* No other underwear does this! Damart Thermolactyl is so comfortable that the Mount Everest climbing expedition wears it. So does the Glencoe mountain rescue team and the entire Chicago Bears Football Club.

Our free color catalog tells the full Damart Thermolactyl story and displays the whole Damart line for men and women. Send for your FREE copy now!

THE PROOF IS IN THE WEARING!

A LIMITED WARRANTY TO CONSUMERS ★ Good Housekeeping PROMISES REPLACEMENT OR REFUND IF DEFECTIVE

Damart Thermawear, Inc.

WHEN IN THE BOSTON AREA, VISIT OUR PORTSMOUTH, N.H. STORE. (603) 431-4700

THERE IS NO WARMER UNDERWEAR MADE!

Fill out and send to:
DAMART, INC. Dept. SN386
1811 Woodbury Ave.
Portsmouth, N.H. 03801

YES! Rush me your **FREE DAMART** Catalog . . . I want to enjoy the fantastic warmth of Thermolactyl Underwear, a DAMART® exclusive. (I understand there is no obligation.)

NAME

ADDRESS

CITY STATE ZIP © 1976, Damart, Inc.

. . . Cryptography

and machine vary in their relative importance. The FBI, which has been involved in cryptanalysis since World War II, is kept quite busy solving criminal ciphers. Since these tend not to be sophisticated, a computer is used infrequently. Most of the codes encountered by the FBI are simple substitution types. Patric W. Paddock, chief of the FBI's cryptanalysis unit, is quick to point out that even when the aid of a computer is summoned, "the [human] cryptanalyst is the one who breaks it, not the machine." The computer, he explains, merely prepares the data and seeks out various repetitive letter patterns for the convenience and decisive scrutiny of the cryptanalyst.

In more sophisticated operations, computers can intervene to a greater extent. They are capable of being programmed to recognize (most, but not all, of the time) text that is syntactically and grammatically correct. As such, machines not only can perform millions of permutations and comparisons, they can "recognize" and record only those (ideally, only one) that are structurally correct. Armed with a reasonably extensive vocabulary, the computer can generate millions of different letter combinations, recalling the one that contains the largest number of sensible words.

Even the most advanced computer systems are vulnerable, however. An extraordinarily complicated code can cause in the computer a "combinatorial explosion." This (usually) figurative expression describes a computer that suddenly finds itself unable to cope with the enormous number of combinations involved in a problem. On the other hand, a computer may be capable of handling the situation without blowing up, but would require longer than the age of the universe.

Besides executing enormously routine operations, the cryptographic computer is prepared to untangle the type of mathematical codes often encountered. For example, a cipher may be based on the mathematical series expansion for the trigonometric quantity, sin 3°. It may be based on the first 26 terms of the solution to the integral problem: $\int x^m(\log x)^n dx$. The possibilities are limitless, and anyone with a mathematical bent is apt to venture into this type of *general system.* All of the FBI's cryptanalysts, says Paddock, are required to have mathematics degrees.

The situation with computers and cryptography is somewhat analogous to the arms race and a treadmill. Although the implements involved reflect a spectacular evolution, the measure of fundamental progress remains unaltered. There are, as always, codes that are not practically decipherable. The horizons of practicality, not the fundamental techniques, in cryptography have evolved along with the modern computer. The state-of-the-art is no longer man versus man, but machine versus machine. □