

The technique used to crack the code is *frequency analysis*: If the cipher is a simple substitution of symbols for letters, then crucial information about which symbols represent which letters can be gathered from how often the various symbols appear in the ciphertext. This idea was first described by the Arabic philosopher and mathematician Al-Kindi, who lived in Baghdad in the ninth century.

By the Renaissance, this kind of informed guesswork had been reduced to a fine art that was well known to European governments. In a famous example of the insecurity of substitution ciphers, Mary Queen of Scots was beheaded in 1587 due to her misplaced reliance on a substitution cipher to conceal her correspondence with plotters against Queen Elizabeth I. She was not the last to have put too much confidence in an encryption scheme that looked hard to crack, but wasn't. Substitution ciphers were in common use as late as the 1800s, even though they had been insecure for a millennium by that time! Edgar Allen Poe's mystery story *The Gold Bug* (1843) and A. Conan Doyle's Sherlock Holmes mystery *Adventure of the Dancing Men* (1903) both turn on the decryption of substitution ciphers.

Secret Keys and One-Time Pads

In cryptography, every advance in code-breaking yields an innovation in code-making. Seeing how easily the *Equatorie* code was broken, what could we do to make it more secure, or *stronger*, as cryptographers would say? We might use more than one symbol to represent the same plaintext letter. A method named for the sixteenth-century French diplomat Blaise de Vigenère uses multiple Caesar ciphers. For example, we can pick twelve Caesar ciphers and use the first cipher for encrypting the 1st, 13th, and 25th letters of the plaintext; the second cipher for encrypting the 2nd, 14th, and 26th plaintext letters; and so on. Figure 5.5 shows such a Vigenère cipher. A plaintext message beginning SECURE... would be encrypted to produce the ciphertext *llqgrw...*, as indicated by the boxed characters in the figure—S is encrypted using the first row, E is encrypted using the second row, and so on. After we use the bottom row of the table, we start again at the top row, and repeat the process over and over.

We can use the cipher of Figure 5.5 without having to send our correspondent the entire table. Scanning down the first column spells out *thomasbryan*, which is the key for the message. To communicate using Vigenère encryption, the correspondents must first agree on a key. They then use the key to construct a substitution table for encrypting and decrypting messages.

When SECURE was encrypted as *llqgrw*, the two occurrences of E at the second and sixth positions in the plaintext were represented by different

ciphertext letters, and the two occurrences of the ciphertext letter *l* represented different plaintext letters. This illustrates how the Vigenère cipher confounds simple frequency analysis, which was the main tool of cryptanalysts at the time. Although the idea may seem simple, the discovery of the Vigenère cipher is regarded as a fundamental advance in cryptography, and the method was considered to be unbreakable for hundreds of years.

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z							
1	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	1						
2	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	2						
3	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	3						
4	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	4						
5	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	5						
6	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	6						
7	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	7						
8	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	8						
9	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	9						
10	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	10						
11	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	11
12	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	12						

Harvard University Archives.

FIGURE 5.5 A Vigenère cipher. The key, *thomasbbryan*, runs down the second column. Each row represents a Caesar cipher in which the shift amount is determined by a letter of the key. (Thomas B. Bryan was an attorney who used this code for communicating with a client, Gordon McKay, in 1894.)

CRYPTOGRAPHY AND HISTORY

Cryptography (code-making) and cryptanalysis (code-breaking) have been at the heart of many momentous events in human history. The intertwined stories of diplomacy, war, and coding technology are told beautifully in two books: *The Code-Breakers*, revised edition, by David Kahn (Scribner's, 1996) and *The Code Book* by Simon Singh (Anchor paperback, 2000).

Cryptographers use stock figures for describing encryption scenarios: Alice wants to send a message to Bob, and Eve is an adversary who may be eavesdropping.

Suppose Alice wants to send Bob a message (see Figure 5.6). The lock-and-key metaphor goes this way: Alice puts the message in a box and locks the box, using a key that only she and Bob possess. (Imagine that the lock on Alice's box is the kind that needs the key to lock it as well as to open it.) If Eve intercepts the