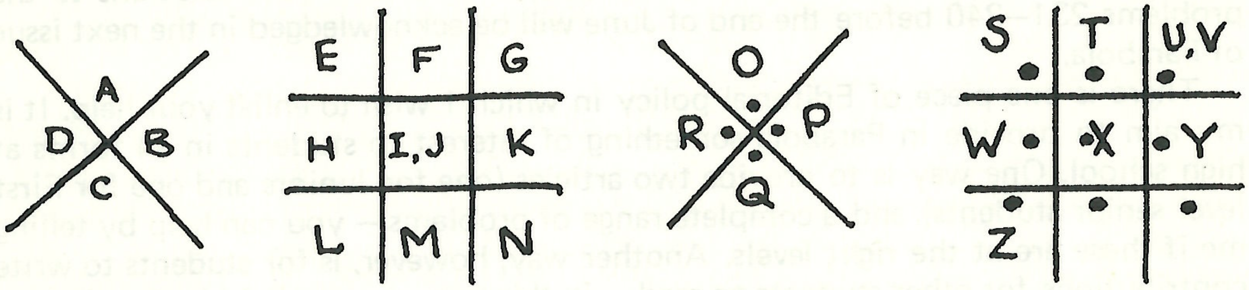
CODES

At some stage of our school life, most of us would have sent or received coded messages from our classmates. A few of us may have even intercepted a coded message and puzzled about its hidden meaning. The coding of messages has been used for many centuries, some codes date back as far as Ancient Greek and Roman times.

Sir Francis Bacon (1561 – 1626) who was a writer, philosopher and Lord Chancellor of the Elizabethan court, gives us three criteria for a good code.

(1) "Easy and not labourious to write."
(2) "Safe and impossible to decipher."
(3) "If possible such as not to raise suspicion."

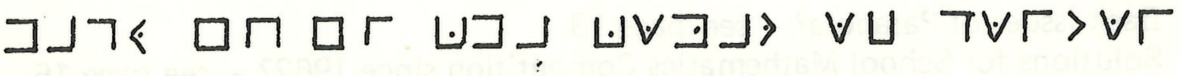
In the sixteenth century a code of quite common usage in England was



Note: The English alphabet at this time had only 24 letters I and J being written the same, and U and V being written the same.

As an example of this code we have:

HELP IM IN THE TOWER OF LONDON



This type of code is called an invented alphabet, as each letter is replaced by a new symbol. There are many ways of producing such codes.

A shifting of the letters of the alphabet by a simple pattern, would be one that you all have thought of:

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

becomes

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

(VMYQE NAZP UE PQMP)

You could also have considered the use of a code word, say MAGAZINE.

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

becomes

M A G Z I N E B C D F H J K L O P Q R S T U V W X Y

(BI LKHX HCUIZ SVCGI).

Note: After the different letters of the code word have been used, the remainder of the alphabet is filled in, in alphabetic order.

Is the invented alphabet code then a good code according to Bacon's criteria?

I think we would all agree that these codes are easily remembered and that there would be no need to keep the code written out on a dangerous piece of paper. Also I think we would all agree that if a coded message of this type were found then it would create a great deal of suspicion to the finder.

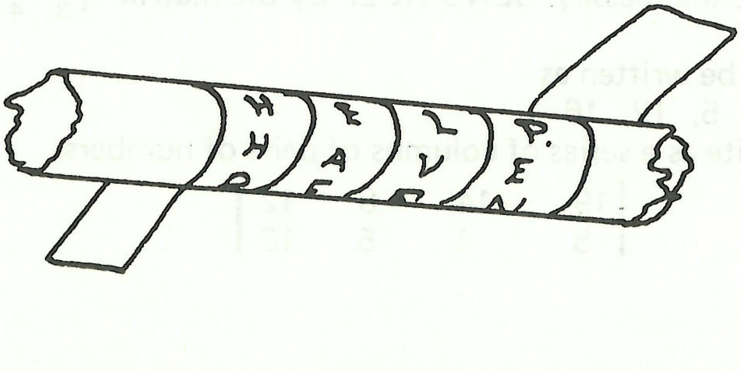
Therefore we are left to answer the question: is the invented alphabet code safe?

At first we may be tempted to answer yes, because if we have 26 different symbols which will replace the 26 letters of our alphabet then there are 26 x 25 x 24 x . . . x 3 x 2 x 1 = 403,291,461,126,605,635,584,000,000 different ways in which we can do this.

However because we require an easily remembered pattern this number would be greatly reduced. Single letters and small words also make our job easier. For example think of all possible decodings for "□ ⊓ □⌈ " from our first example. You should be able to conclude that "□" is the replacement of an A or an I.

It is true that in our examples so far we have been more than generous to the code breaker by leaving, unnecessarily, the spaces between the words. By removing these spaces his task is made harder but he still has methods to use if sufficient coded messages are intercepted. It is quite commonly known that the letter "e" is used more frequently than any other letter in English. Some letter pairs occur quite often (th, ed,). This sort of information applied to enough messages will eventually break invented alphabet codes. We ought to conclude therefore that invented alphabets are not safe, and therefore not good codes.

GEOMETRIC CODES. Another type of code is based on geometric patterns. The Romans used a strip of paper, which revealed its message when it was wound around a particular sized cylinder, say a spear handle.



Other codes of this type would be:

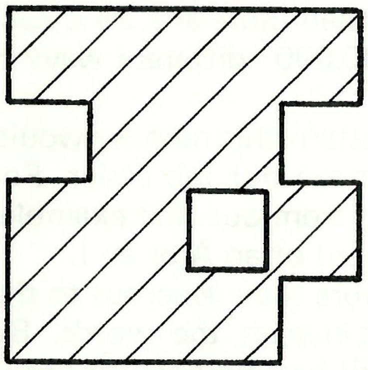
```
P A R A
B O L A
I S A M   transmit:
A T H E   PBIAMCGNAOSTASAERLAHTMZPAAMEIAIQ.
M A T I
C S M A   OR
G A Z I   PARAAMEIAIQPENGCMAIBOLAHTMZASATS.
N E P Q
```

Work out the patterns that have been used.

Any other pattern of letters would be just as acceptable as the above two.

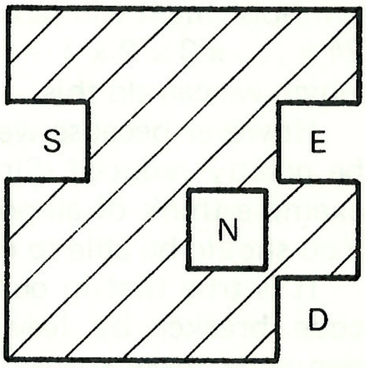
Another device which uses a geometric pattern is the grille. This is a square with a set of holes cut in it, in such a way that all the letters of the message are revealed by rotating the grille 3 times through 90° in a clockwise direction.

e.g. GRILLE



MESSAGE

M	Y	M	N
S	O	O	E
N	O	N	E
R	W	E	D



Copy the above grille onto a piece of paper, cut it out, and use it to read the message (the first word is done for you).

MATRIX CODES. Another method of coding uses matrices. A matrix is a rectangular pattern of numbers. Let us begin by numbering the letters of the alphabet.

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	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

Now we will code the message SEND HELP by the matrix $\begin{bmatrix} 4 & 5 \\ 3 & 4 \end{bmatrix}$.

SEND HELP can be written as
19, 5, 14, 4, 8, 5, 12, 16
which we will write as a series of columns of pairs of numbers:

$$\begin{bmatrix} 19 & 14 & 8 & 12 \\ 5 & 4 & 5 & 16 \end{bmatrix}$$

Now we apply the code matrix to each of these columns as follows:

$$\begin{bmatrix}4&5\\3&4\end{bmatrix}\begin{bmatrix}19\\5\end{bmatrix} = \begin{bmatrix}4\times19+5\times5\\3\times19+4\times5\end{bmatrix} = \begin{bmatrix}101\\77\end{bmatrix}$$

You should notice carefully how each number in the result is formed. The first number is a combination of the numbers from the top row of the coding matrix with the numbers 19 and 5, and so on.

Repeating this process for each column, we can code the whole message:

$$\begin{bmatrix}4&5\\3&4\end{bmatrix}\begin{bmatrix}19&14&8&12\\5&4&5&16\end{bmatrix}=\begin{bmatrix}101&76&57&128\\77&58&44&100\end{bmatrix}$$

At this stage, as many multiples of 26 are substracted as is possible: only the remainders are left in the corresponding position.

i.e.

$$\begin{bmatrix}23&24&5&24\\25&6&18&22\end{bmatrix}$$

which could be written as 23, 25, 24, 6, 5, 18, 24, 22 or sent as WYXF ERXV.

You will notice that the E in SEND has been coded as Y while the E in HELP has been coded as R. Also in the coded message, the X is the code of N and L. So the decoding procedures for the invented alphabet codes would not work on matrix coding.

To decode the message the procedure is precisely the same as for coding the message, except that we use the decoding matrix $\begin{bmatrix}4&-5\\-3&4\end{bmatrix}$.

Thus WYXF→23, 25, 24, 6.

Decoding,

$$\begin{bmatrix}4&-5\\-3&4\end{bmatrix}\begin{bmatrix}23\\25\end{bmatrix} = \begin{bmatrix}4\times23-5\times25\\-3\times23+4\times25\end{bmatrix} = \begin{bmatrix}-33\\31\end{bmatrix}$$

$$\rightarrow \begin{bmatrix}19\\5\end{bmatrix} \rightarrow \text{SE}$$

and $\begin{bmatrix}4&-5\\-3&4\end{bmatrix}\begin{bmatrix}24\\6\end{bmatrix} = \begin{bmatrix}66\\-48\end{bmatrix}\rightarrow\begin{bmatrix}14\\4\end{bmatrix} \rightarrow \text{ND}$

You will notice that, when the entries were negative, 26's were added on until the remainder was first positive.

There is a simple rule to tell you the decoding matrix from the encoding matrix. If the encoding matrix is $\begin{bmatrix}a&b\\c&d\end{bmatrix}$ then the decoding matrix is

$$\begin{pmatrix}\frac{d}{ad-bc}&\frac{-b}{ad-bc}\\[1ex]\frac{-c}{ad-bc}&\frac{a}{ad-bc}\end{pmatrix}$$

You might like to decode ERXV and see if you get it right.

All the codes we have seen so far raise suspicion when discovered and so would not meet Bacon's third criterion for a good code. But perhaps this is not as important as it used to be, because in Bacon's day codes were carried by people and passed on. These days most messages would be sent by wireless and, although more liable to be intercepted, do not produce the same dangers to the senders and receivers if they cannot be broken.

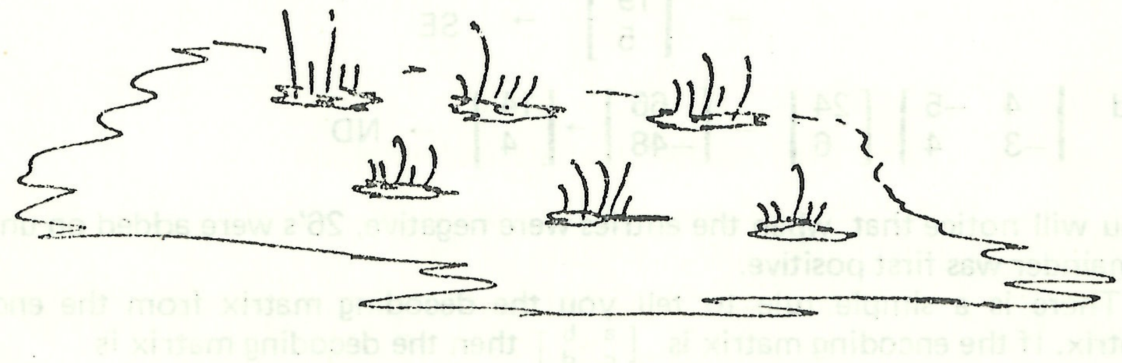
Bacon is thought to be responsible for a code that might not raise suspicion if, by chance, it was intercepted. The code is disguised in a normal letter and works on the binary numbering system using groups of five letters.

A = 00001, B = 00010, C = 00011, D = 00100, ...

I am quite sure that you are familiar with the binary numbering system and can finish the coding of this alphabet. If the capitals are replaced by one's and small letters are replaced by zeros, try to decode the following ransom note:

dEar mum AnD dAD,
the kIdnappeRS wAnt one
ThouSaNd dolLars FOr My safE retuRN.
YOUr LoVIng SOn,
bILL.

During the Second World War, the British Intelligence were sure that a person was sending information to Germany. He was watched very closely and they had found that he sent paintings to Germany by various devious routes. They were convinced that the paintings held the messages. After many puzzling hours they found that the blades of grass were painted in a binary code.



(Answers to the codes are given on page 19.)

ACTIVITIES:

(a) Select a page in a novel and calculate the percentage of times the letter "e" is used on that page.

(b) Code a message with the matrix $\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$. What happens?

(c) Decode a coded message whose matrix is the same as the coding matrix. What do you find? Does this always happen?

(d) Find the decoding matrix for $\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}$. What other matrices cannot be used for coding?

☆ ☆ ☆

Solution to Crossnumber "1" of Vol. 10 No. 1

a 1	b 8	c 6	0	d 8	e 6	f 7
g 5	7	4	■	h 5	1	2
i 1	6	■	■	j 1	2	3
k 5	5	m 0	■	n 8	9	1
o 1	2	1	■	p 4	6	2
q 5	1	■	■	■	r 1	7
s 1	0	2	0	1	2	1

Crossnumber 1 — Successful Solvers

Alan Fekete, Sydney Grammar.
Lee Lawrence, Morisset High.
Vincent Quigley, Marist Brothers, Pagewood.
Scott Marshall, James Ruse Ag. High