Simon Singh *The Code Book. The Evolution of Secrecy from Mary Queen of Scots to Quantum Cryptography*
DOUBLEDAY, a division of Random House, Inc.

In the following example, I have enciphered a piece of ciphertext using the Vigenere cipher, using a keyphrase that is as long as the message. All the cryptanalytic techniques that I have previously described will fail. None the less, the message can be deciphered.

Key	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Plaintext	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Ciphertext	**V**	**H**	**R**	**M**	**H**	**E**	**U**	**Z**	**N**	**F**	**Q**	**D**	**E**	**Z**	**R**	**W**	**X**	**F**	**I**	**D**	**K**

This new system of cryptanalysis begins with the assumption that the ciphertext contains some common words, such as **the**. Next, we randomly place **the** at various points in the plaintext, as shown below, and deduce what sort of keyletters would be required to turn **the** into the appropriate ciphertext. For example, if we pretend that **the** is the first word of the plaintext, then what would this imply for the first three letters of the key? The first letter of the key would encrypt **t** into **V**. To work out the first letter of the key, we take a Vigenere square, look down the column headed by **t** until we reach **V**, and find that the letter that begins that row is **C**. This process is repeated with **h** and **e**, which would be encrypted as **H** and **R** respectively, and eventually we have candidates for the first three letters of the key, **CAN**. All of this comes from the assumption that **the** is the first word of the plaintext. We place **the** in a few other positions, and, once again, deduce the corresponding keyletters. (You can check the relationship between each plaintext letter and ciphertext letter by referring to the Vigenere square in Table 9.)

Key	**C**	**A**	**N**	?	?	?	**B**	**S**	**J**	?	?	?	?	?	**Y**	**P**	**T**	?	?	?	?
Plaintext	**t**	**h**	**e**	?	?	?	**t**	**h**	**e**	?	?	?	?	?	**t**	**h**	**e**	?	?	?	?
Ciphertext	**V**	**H**	**R**	**M**	**H**	**E**	**U**	**Z**	**N**	**F**	**Q**	**D**	**E**	**Z**	**R**	**W**	**X**	**F**	**I**	**D**	**K**

We have tested three **the**'s against three arbitrary fragments of the ciphertext, and generated three guesses as to the elements of certain parts of the key. How can we tell whether any of the **the**'s are in the right position? We suspect that the key consists of sensible words, and we can use this to our advantage. If a **the** is in a wrong position, it will probably result in a random selection of keyletters. However, if it is in a correct position, the keyletters should make some sense. For example, the first **the** yields the keyletters **CAN**, which is encouraging because this is a perfectly reasonable English syllable. It is possible that this **the** is in the correct position. The second **the** yields **BSJ**, which is a very peculiar combination of consonants, suggesting that the second **the** is probably a mistake. The third **the** yields **YPT**, an unusual syllable but one which is worth further investigation. If **YPT** really were part of the key, it would be within a larger word, the only possibilities being **APOCALYPTIC, CRYPT** and **EGYPT,** and derivatives of these words. How can we find out if one of these words is part of the key? We can test each hypothesis by inserting the three candidate words in the key, above the appropriate section of the ciphertext, and working out the corresponding plaintext:

Key	**C**	**A**	**N**	?	?	?	?	**?**	**A**	**P**	**O**	**C**	**A**	**L**	**Y**	**P**	**T**	**I**	**C**	?	?
Plaintext	**t**	**h**	**e**	?	?	?			**n**	**q**	**c**	**b**	**e**	**o**	**t**	**h**	**e**	**x**	**g**	?	?
Ciphertext	**V**	**H**	**R**	**M**	**H**	**E**	**U**	**Z**	**N**	**F**	**Q**	**D**	**E**	**Z**	**R**	**W**	**X**	**F**	**I**	**D**	**K**

Key	**C**	**A**	**N**	?	?	?	?	?	?	?	?	?	**C**	**R**	**Y**	**P**	**T**	?	?	?	?
Plaintext	**t**	**h**	**e**	?	?	?	?	?	?	?	?	?	**c**	**i**	**t**	**h**	**e**	?	?	?	?
Ciphertext	**V**	**H**	**R**	**M**	**H**	**E**	**U**	**Z**	**N**	**F**	**Q**	**D**	**E**	**Z**	**R**	**W**	**X**	**F**	**I**	**D**	**K**

Key	**C**	**A**	**N**	?	?	?	?	?	?	?	?	?	**E**	**G**	**Y**	**P**	**T**	?	?	?	?
Plaintext	**t**	**h**	**e**	?	?	?	?	?	?	?	?	?	**a**	**t**	**t**	**h**	**e**	?	?	?	?
Ciphertext	**V**	**H**	**R**	**M**	**H**	**E**	**U**	**Z**	**N**	**F**	**Q**	**D**	**E**	**Z**	**R**	**W**	**X**	**F**	**I**	**D**	**K**

If the candidate word is not part of the key, it will probably result in a random piece of plaintext, but if it is part of the key the resulting plaintext should make some sense. With **APOCALYPTIC** as part of the key the resulting plaintext is gibberish of the highest quality. With **CRYPT,** the resulting plaintext is **cithe**, which is not an inconceivable piece of plaintext. However, if **EGYPT** were part of the key it would generate **atthe**, a more promising combination of letters, probably representing the words **at the**.

For the time being let us assume that the most likely possibility is that **EGYPT** is part of the key. Perhaps the key is a list of countries. This would suggest that **CAN**, the piece of the key that corresponds to the first **the**, is the start of **CANADA.** We can test this hypothesis by working out more of the plaintext, based on the assumption that **CANADA,** as well as **EGYPT,** is part of the key:

Key	**C**	**A**	**N**	**A**	**D**	**A**	?	?	?	?	?	?	**E**	**G**	**Y**	**P**	**T**	?	?	?	?
Plaintext	**t**	**h**	**e**	**m**	**e**	**e**	?	?	?	?	?	?	**a**	**t**	**t**	**h**	**e**	?	?	?	?
Ciphertext	**V**	**H**	**R**	**M**	**H**	**E**	**U**	**Z**	**N**	**F**	**Q**	**D**	**E**	**Z**	**R**	**W**	**X**	**F**	**I**	**D**	**K**

Our assumption seems to be making sense. **CANADA** implies that the plaintext begins with **themee** which perhaps is the start of **the meeting.** Now that we have deduced some more letters of the plaintext, **ting,** we can deduce the corresponding part of the key, which turns out to be **BRAZ.** Surely this is the beginning of **BRAZIL.** Using the combination of **CANADABRAZILEGYPT** as the bulk of the key, we get the following decipherment: **the meeting is at the ????.**

In order to find the final word of the plaintext, the location of the meeting, the best strategy would be to complete the key by testing one by one the names of all possible countries, and deducing the resulting plaintext. The only sensible plaintext is derived if the final piece of the key is **CUBA:**

Key	**C**	**A**	**N**	**A**	**D**	**A**	**B**	**R**	**A**	**Z**	**I**	**L**	**E**	**G**	**Y**	**P**	**T**	**C**	**U**	**B**	**A**
Plaintext	**t**	**h**	**e**	**m**	**e**	**e**	**t**	**i**	**n**	**g**	**i**	**s**	**a**	**t**	**t**	**h**	**e**	**d**	**o**	**c**	**k**
Ciphertext	**V**	**H**	**R**	**M**	**H**	**E**	**U**	**Z**	**N**	**F**	**Q**	**D**	**E**	**Z**	**R**	**W**	**X**	**F**	**I**	**D**	**K**

Table 9 Vigenere square.

Plain	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	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
2	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	B
3	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
4	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
5	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
6	G	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
7	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
8	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	H
9	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
10	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
11	L	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
12	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
13	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
14	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
15	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	O
16	Q	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
17	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
18	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
19	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
20	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
21	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
22	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
23	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
24	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
25	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
26	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

So, a key that is as long as the message is not sufficient to guarantee security. The insecurity in the example above arises because the key was constructed from meaningful words. We began by randomly inserting **the** throughout the plaintext, and working out the corresponding keyletters. We could tell when we had put a **the** in the correct place, because the keyletters looked as if they might be part of meaningful words. Thereafter, we used these snippets in the key to deduce whole words in the key. In turn this gave us more snippets in the message, which we could expand into whole words, and so on. This entire process of toing and froing between the message and the key was only possible because the key had an inherent structure and consisted of recognisable words. However, in 1918 cryptographers began experimenting with keys that were devoid of structure. The result was an unbreakable cipher.