

Episode 12 – Breaking a substitution code

European section – Season 3

A Manuscript on Deciphering Cryptographic Messages.

by Abu Yusuf Ya 'qub ibn Is-haq ibn as-Sabbah
ibn 'omran ibn Ismail **al-Kindi** (9th century).

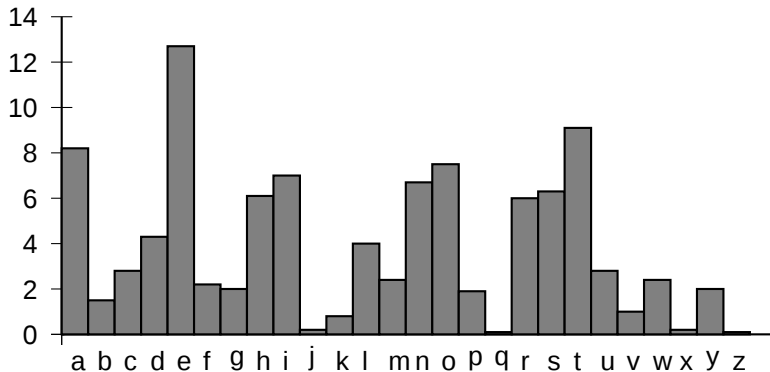
[illegible]

الحمد لله - والحمد لله، والاعمال الصالحة علم به محمد وآل بيته ع

السلام عليكم ورحمة الله وبركاته
 والحمد لله رب العالمين
 وبعد
 أما بعد
 فإني أبلغكم بهذا الكتاب
 الذي كتبه لي
 في تاريخ
 الإسلام
 من سنة
 ١٢٠٠
 إلى سنة
 ١٢٨٠
 من الهجرة
 النبوية
 صلى الله عليه وسلم
 وآله
 وسلم
 وأما
 بعد
 فإني أبلغكم
 بهذا الكتاب
 الذي كتبه لي
 في تاريخ
 الإسلام
 من سنة
 ١٢٠٠
 إلى سنة
 ١٢٨٠
 من الهجرة
 النبوية
 صلى الله عليه وسلم
 وآله
 وسلم

Frequency analysis

Letter frequencies in English



The Gold Bug cipher

53†††305))6*;4826)4†.)4†);806*;48†8
\$60))85;1†(;:†*8†83(88)5*†;46(;88*96
?;8)†(;485);5*†2:*†(;4956*2(5*-4)8
\$8*;4069285);)6†8)4††;1(†9;48081;8:8†
1;48†85;4)485†528806*81(†9;48;(88;4
(†?34;48)4†;161;:188;†?;

The Gold Bug cipher : A solution

Symbol frequencies in the cipher

8	;	4	‡	)	*	5	6	†	1
33	26	19	16	16	13	12	11	8	8

0	9	2	:	3	?	\$	—	.
6	5	5	4	4	3	2	1	1

The Gold Bug cipher : A solution

$$8 = e$$

The Gold Bug cipher : A solution

$$\frac{8}{e}$$

The Gold Bug cipher : A solution

$$\frac{8}{e}$$

;48 =

The Gold Bug cipher : A solution

$$\frac{8}{e}$$

;48 = the

The Gold Bug cipher : A solution

8	;	4
e	t	h

The Gold Bug cipher : A solution

8	;	4
e	t	h

;48;(88;4 =

The Gold Bug cipher : A solution

8	;	4
e	t	h

;48;(88;4 = thet.eeth =

The Gold Bug cipher : A solution

8		;		4
e		t		h

;48;(88;4 = thet.eeth = the tree th

The Gold Bug cipher : A solution

8	;	4	(
e	t	h	r

The Gold Bug cipher : A solution

8	;	4	(
e	t	h	r

the tree ;4(†?34

=

The Gold Bug cipher : A solution

8	;	4	(
e	t	h	r

the tree ;4(†?34
= the tree thr†?3h
=

The Gold Bug cipher : A solution

8	;	4	(
e	t	h	r

the tree ;4(†?34
= the tree thr†?3h
= the tree through

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3
e	t	h	r	o	u	g

The Gold Bug cipher : A solution

8	;	4	(	†	?	3
e	t	h	r	o	u	g

†83(88 =

The Gold Bug cipher : A solution

8	;	4	(	†	?	3
e	t	h	r	o	u	g

†83(88 = †egree =

The Gold Bug cipher : A solution

8	;	4	(	†	?	3
e	t	h	r	o	u	g

†83(88 = †egree = degree

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†
e	t	h	r	o	u	g	d

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†
e	t	h	r	o	u	g	d

;46(;88*=

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†
e	t	h	r	o	u	g	d

;46(;88*= th6rtee*=

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†
e	t	h	r	o	u	g	d

;46(;88*= th6rtee*= thirteen

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†	6	*
e	t	h	r	o	u	g	d	i	n

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†	6	*
e	t	h	r	o	u	g	d	i	n

53‡‡‡=

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†	6	*
e	t	h	r	o	u	g	d	i	n

53‡‡‡† = 5good =

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†	6	*
e	t	h	r	o	u	g	d	i	n

53‡‡‡† = 5good = A good

The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†	6	*	5
e	t	h	r	o	u	g	d	i	n	a

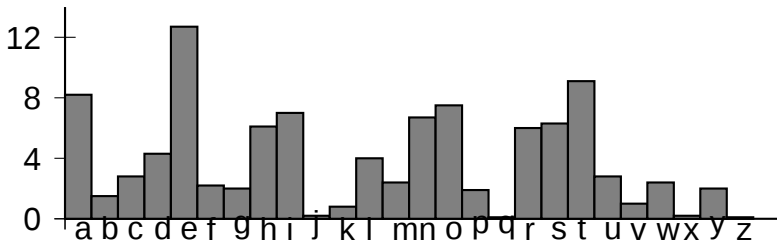
The Gold Bug cipher : A solution

8	;	4	(	‡	?	3	†	6	*	5
e	t	h	r	o	u	g	d	i	n	a

'A good glass in the bishop's hostel in the devil's seat forty-one degrees and thirteen minutes northeast and by north main branch seventh limb east side shoot from the left eye of the death's-head a bee-line from the tree through the shot fifty feet out.'

Tips for frequency analysis

- Count up the frequencies of all the letters in the ciphertext.
- About five of the letters should have a frequency less than 1 per cent, they probably represent j, k, q, x and z.
- One of the letters should have a frequency greater than 10 percent, and it probably represents e.



Tips for frequency analysis

- If the plaintext doesn't reveal itself immediately, which is often the case, then focus on pairs of repeated letters.
- In English the most common repeated letters are ss, ee, tt, ff, ll, mm and oo.
- If the ciphertext contains any repeated characters, you can assume that they represent one of these.

Tips for frequency analysis

- More complex use of statistics can be conceived, such as considering counts of pairs of letters, or triplets (trigrams), and so on.
- For instance, q and u nearly always occur together in that order in English, even though q itself is rare.
- Also, *th* is the most common bigram, and *the* the most common trigram.

Tips for frequency analysis

- One of the most useful skills for a cryptanalyst is the ability to identify words, or even entire phrases, based on experience or sheer guesswork. Any such word or phrase is known as a *crib*.

Cipher 8



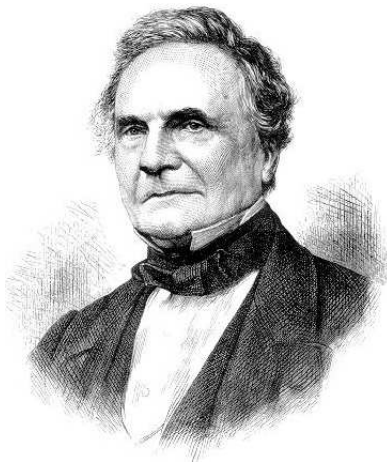
Al-Kindi

Cipher 6



Blaise de Vigenère

Cipher 1



Charles Babbage

Cipher 2



Alan Turing

Cipher 5

Cipher 5



Ralph Merkle, Martin Hellman, Whitfield Diffie

Cipher 4

Cipher 4



Adi Shamir, Ronald Rivest, Len Adleman

Cipher 3



Phil Zimmermann

Cipher 7



Clifford Cocks

Cipher 9



Georges Perec