Breaking a Transposition Cipher

Say we have some ciphertext that we know was encrypted with a transposition cipher. At first, we assume we know the degree of the permutation.

Say the degree is 13. We arrange our ciphertext into 13 columns (perhaps disregarding an incomplete last row).

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
t	**i**	**f**	**a**	**t**	**p**	**o**	**k**	**g**	**r**	**i**	**a**	**n**
e	**b**	**s**	**t**	**m**	**n**	**e**	**l**	**r**	**t**	**i**	**a**	**e**
t	**c**	**t**	**n**	**i**	**s**	**h**	**s**	**e**	**s**	**i**	**n**	**i**
d	**f**	**e**	**n**	**a**	**h**	**m**	**e**	**u**	**s**	**t**	**v**	**o**
o	**e**	**e**	**m**	**a**	**t**	**r**	**l**	**t**	**a**	**s**	**l**	**s**
l	**i**	**e**	**p**	**t**	**n**	**y**	**l**	**e**	**t**	**a**	**h**	**k**
s	**s**	**r**	**g**	**e**	**a**	**t**	**u**	**n**	**s**	**t**	**t**	**i**
h	**o**	**o**	**t**	**o**	**p**	**e**	**o**	**t**	**w**	**e**	**b**	**s**
r	**t**	**l**	**i**	**r**	**e**	**s**	**f**	**e**	**e**	**t**	**o**	**h**
e	**e**	**w**	**h**	**c**	**e**	**l**	**p**	**t**	**p**	**n**	**o**	**s**
b	**t**	**l**	**a**	**a**	**l**	**f**	**e**	**s**	**o**	**a**	**n**	**i**
t	**k**	**i**	**a**	**r**	**d**	**f**	**c**	**o**	**n**	**e**	**o**	**c**
e	**e**	**s**	**c**	**r**	**n**	**o**	**z**	**o**	**r**	**a**	**o**	**n**
v	**l**	**i**	**i**	**u**	**t**	**e**	**h**	**f**	**m**	**e**	**g**	**y**
e	**e**	**e**	**z**	**s**	**f**	**r**	**h**	**i**	**a**	**o**	**t**	**s**
o	**t**	**t**	**m**	**t**	**e**	**a**	**i**	**s**	**h**	**r**	**h**	**i**
n	**o**	**i**	**a**	**n**	**l**	**h**	**h**	**s**	**g**	**d**	**t**	**u**
w	**a**	**i**	**a**	**t**	**y**	**p**	**e**	**s**	**m**	**a**	**y**	**r**
o	**i**	**o**	**y**	**o**	**t**	**r**	**s**	**l**	**f**	**u**	**s**	**b**
f	**h**	**e**	**r**	**e**	**r**	**e**	**c**	**o**	**s**	**o**	**a**	**f**
m	**l**	**t**	**u**	**e**	**n**	**a**	**b**	**h**	**r**	**a**	**a**	**e**
h	**b**	**a**	**t**	**v**	**m**	**n**	**l**	**t**	**n**	**e**	**e**	**u**
d	**f**	**e**	**s**	**e**	**t**	**t**	**o**	**t**	**c**	**e**	**r**	**i**
e	**s**	**s**	**g**	**i**	**t**	**h**	**n**	**g**	**t**	**s**	**o**	**u**
s	**i**	**o**	**r**	**e**	**c**	**l**	**n**	**e**	**u**	**e**	**u**	**v**
o	**f**	**i**	**l**	**d**	**m**	**s**	**e**	**l**	**d**	**f**	**b**	**u**
r	**d**	**t**	**r**	**i**	**s**	**i**	**e**	**e**	**r**	**e**	**z**	**t**
t	**l**	**o**	**e**	**a**	**n**	**t**	**p**	**n**	**t**	**s**	**l**	**a**
r	**i**	**t**	**e**	**o**	**n**	**r**	**d**	**f**	**f**	**e**	**o**	**f**
n	**u**	**t**	**w**	**o**	**e**	**i**	**o**	**o**	**h**	**w**	**m**	**r**

The frequencies of individual characters, by themselves, don't help us. They are the same as in the plaintext, and don't depend on the key (the permutation used to rearrange the columns).

But the frequencies of digrams can be very helpful.

Consider just 4 characters from the cibertext above.

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
row 4								**e**			**t**		
row 12		**k**						**c**					

Based solely on the information above, which column, 2 or 11, is more likely to come immediately after column 8 in the plaintext? In other words, which is more likely in the plaintext?

		8	*2*	
row 4	**...**	**e**		
row 12	**...**	**c**	**k**	**.**

		8	*11*	
row 4	**...**	**e**	**t**	**...**
row 12	**...**	**c**		**...**

i) In the first case (columns 8,2), the plaintext has a digram **ck**. We estimated $prob(\mathbf{ck}) = 10/10000$.

ii) In the second case (columns 8,11), the plain text has a digram **et**. We estimated $prob(\mathbf{et}) = 83/10000$.

Since **et** occurs about eight times as often as **ck**, in English text, we might conclude the second case (8,11) is more likely to occur, based on our limited information.

But consider more carefully. Suppose column 8 and 11 are “far” apart in the plaintext. The the letters in columns 8 and 11 of some row are nearly independent. This means that

$$prob(\mathbf{et}) = prob(\mathbf{e})\,prob(\mathbf{t}) = (1237/10000)(921/10000) = 114/10000.$$

Thus the digram **et** is more likely to occur in plaintext columns that are well separated, than in adjacent columns ($114/83 \approx 1.37$ times as likely).

In other words, **et** is a common digram only because **e** and **t** are common letters. Given the frequencies of **e** and **t**, the frequency of **et** is lower than one would expect.

So the **et** in columns 8,11 makes it less likely that column 11 comes immediately after column 8 in the plaintext.

Now consider the **ck** in columns 8,2. If columns 8 and 2 are well separated in the plaintext (so their contents in a given row are nearly independent), then

$$prob(\mathbf{ck}) = prob(\mathbf{c})\,prob(\mathbf{k}) = (230/10000)(87/10000) = 2/10000.$$

Thus **ck** is five times more likely to occur in two adjacent plaintext columns, than in two columns that are well separated. Although **ck** is a fairly rare digram, it is far more common than would expect, based on the low frequencies of **c** and (expecially) **k**.

Thus a **ck** in columns 8,2 makes it considerably more likely that column 2 comes immediately after column 8.

Naturally, to make a better guess, we should look at all the digrams in columns 8,11 and columns 8,2.

More generally, consider two columns, *i* and *j*, and consider a pair λμ of characters. By *prob*(λμ) we mean the probability that λ occurs in column *i* and μ in column *j* (of the same row).

$$\text{Then } prob(\lambda\mu) = \begin{cases} prob(\lambda)prob(\mu) & \text{if rows } i \text{ and } j \text{ are well separated, in the plaintext,} \\ prob(\text{digram } \lambda\mu \text{ in Engl lang text}) & \text{if row } j \text{ immediately follows row } i \text{ in the plaintext.} \end{cases}$$

Of course, not all cases are covered above. For example, column *j* might comes two columns after column *i*, in the plaintext. If this occurs, the characters in columns *i* and *j* of a row are not independent, but they are closer to independent than if column *j* comes directly after column *i* in the plaintext (and the nature of the dependency is different). For simplicity, we shall treat these columns as independent, recognizing this will introduce some error.

For each digram λμ, we compute $\boldsymbol{prob(\lambda\mu)/(prob(\lambda)prob(\mu))}$.

i) If $prob(\lambda\mu)/(prob(\lambda)prob(\mu)) > 1$, λμ in columns *i,j* makes it more likely that column *j* is the column that follows column *i* in the plaintext. (The larger $prob(\lambda\mu)/(prob(\lambda)prob(\mu))$ is, the stronger the effect.)

ii) If $prob(\lambda\mu)/(prob(\lambda)prob(\mu)) < 1$, λμ in columns *i,j* makes it less likely that column *j* is the column that follows column *i* in the plaintext.

In the table on the following page, I have given each of the $26^2 = 676$ digrams λμ a score between -8 and 8, with higher scores indicating a higher values of $prob(\lambda\mu)/(prob(\lambda)prob(\mu))$.

A score of 0 indicates $prob(\lambda\mu)/(prob(\lambda)prob(\mu)) \approx 1$ (specifically, between 1/1.189 and 1.189 .

A score of 8 indicates $prob(\lambda\mu)/(prob(\lambda)prob(\mu)) > 13.45$, and -8 indicates $prob(\lambda\mu)/(prob(\lambda)prob(\mu)) < 1/13.45$.

I may explain in class more about how the scores were assigned.

For a pair of *i, j* of columns, we can easily compute the score of each row in columns *i, j* and then compute the average score (over all rows) for columns *i* and *j*.

i) If column *j* immediately follows column *i* in the plaintext, we expect this average score to be roughly

$$\sum_{\text{all di grams } \lambda\mu} prob(\lambda\mu)\, score(\lambda\mu) \approx 1.15.$$

	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
a	-8	1	2	1	-8	-1	0	-7	-1	-2	2	2	1	3	-8	1	-1	1	2	1	-3	3	-1	-2	1	2
b	0	-2	-8	-8	3	-8	-8	-8	-1	3	-8	3	-6	-8	1	-8	-8	1	-4	-7	5	-8	-8	-8	4	-8
c	1	-8	-1	-8	1	-8	-8	3	-1	-8	6	0	-8	-8	3	-8	3	0	-8	-1	0	-8	-8	-8	-4	-2
d	1	3	0	-2	0	0	-1	-1	1	3	-3	-1	0	-2	0	-1	0	-2	0	1	-1	-2	1	-8	0	-8
e	0	1	1	2	-3	0	-2	-3	-2	0	-2	0	1	1	-3	1	2	3	1	-1	-5	2	1	5	0	-1
f	1	-1	-2	-5	-1	2	-4	-2	1	1	-7	0	0	-8	2	-1	-3	1	-3	2	1	-4	-1	-8	-2	-6
g	1	0	-3	-5	0	-1	0	3	1	1	-7	0	-2	-4	1	-3	-3	0	-1	-1	1	-4	0	-8	-4	-8
h	2	-5	-7	-8	4	-6	-8	-5	2	-3	-8	-8	-5	-8	0	-6	-5	-4	-6	-2	-2	-8	-4	-8	-3	-8
i	-5	-2	2	0	-4	0	2	-6	-8	-1	1	1	2	4	-2	-2	-2	-1	2	1	-8	3	-2	2	-8	5
j	1	-8	-8	-8	0	-8	-8	-8	2	-5	-8	-8	-8	-8	2	-8	-8	-8	-8	-8	8	-8	-8	-8	-8	-8
k	0	0	-4	-6	3	-1	-6	-3	3	0	-8	-2	-3	1	-1	-4	-2	-7	0	-2	-3	-7	0	-8	0	-8
l	1	-2	-4	2	1	0	-5	-5	2	-3	0	4	-3	-7	1	-2	-3	-6	-2	-3	-1	-1	-2	-8	4	1
m	2	2	-5	-8	2	-3	-6	-5	1	-1	-7	-6	-1	-7	1	3	-4	-4	-1	-3	0	-8	-2	-8	4	-8
n	-1	-1	1	4	-1	-2	5	-3	-1	1	1	-3	-3	-5	0	-3	1	-8	-1	1	-4	-3	-1	-3	-1	-3
o	-4	0	-1	-2	-8	4	-2	-4	-4	-1	2	0	3	2	-1	1	-2	2	-2	-1	5	2	3	-2	-3	0
p	2	-4	-7	-8	1	-5	-8	-3	0	-3	-8	3	-5	-8	2	4	-4	2	-2	-1	1	-8	-3	-8	-3	-8
q	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	8	-8	-8	-8	-8	-8
r	0	-1	-1	-1	2	-1	0	-3	1	0	1	-2	0	-2	1	-1	0	-2	0	-1	-1	0	-1	-3	2	-8
s	1	1	0	-4	0	-1	-3	1	0	1	0	-2	-1	-4	1	2	3	-6	0	2	0	-3	1	-8	-3	0
t	-1	-1	-2	-5	-1	-3	-5	5	1	0	-4	-2	-2	-6	1	-4	-2	-2	-2	-1	-1	-7	1	-8	0	-5
u	-4	0	2	-2	-5	-3	2	-7	-3	-3	-2	3	0	2	-8	4	-6	3	2	2	-8	-5	-3	-4	-7	2
v	-1	-8	-8	-8	5	-8	-8	-8	2	-8	-8	-8	-8	-8	-1	-8	-8	-8	-8	-8	-8	-8	-8	-8	-3	-8
w	3	-4	-5	-6	1	-5	-8	3	2	-2	-6	-5	-4	-1	1	-6	-6	-5	-4	-5	-7	-7	-3	-8	-4	-4
x	0	-4	5	-7	-1	-3	-8	-2	2	-8	-8	-8	-3	-8	-5	7	4	-8	-5	2	-3	0	-3	6	-2	-8
y	0	3	1	-1	-2	1	-1	-1	0	3	-2	-2	1	-4	3	0	1	-3	1	0	-4	-1	2	-8	-2	-8
z	1	4	-7	-7	4	-7	-5	-6	1	-4	0	0	-5	-8	-1	-5	-8	-8	-8	-7	-6	-8	-8	-8	2	8

The entry in row λ, column μ is the score for digram λμ.

ii) Otherwise, we expect the average score to be roughly

$$\sum_{\text{all di grams } \lambda\mu} prob(\lambda)\, prob(\mu)\, score(\lambda\mu) \;\approx\; \text{-}1.41,$$

although this may be somewhat off if column *j* lies close to column *i* in the plaintext.

If we compute the average score for every pair of columns in the ciphertext given earlier, we obtain the following matrix. The entry in row *i*, column *j* is the score for column pair *i*, *j*.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1		-1.7	-0.6	-0.8	-1.1	-2.1	-0.6	-1.4	-1.4	-1.6	**1.2**	-1.9	-1.5
2	-1.5		-0.8	-2.8	-1.1	-1.6	-0.8	-1.5	-1.6	-1.8	-1.5	-1.3	**1.1**
3	-1.0	-1.2		-1.0	-1.1	-0.7	-0.4	-1.7	-1.3	**0.7**	-1.6	-0.8	-1.9
4	**1.1**	-0.6	-0.7		-1.1	-1.0	-1.7	-1.4	-0.5	-0.6	-1.6	-2.2	-1.8
5	-0.7	-0.3	-1.6	-1.2		-0.9	-1.8	-2.3	-1.1	-0.7	-1.2	**1.5**	-2.5
6	-2.1	-1.5	**0.5**	-1.6	-1.4		-0.7	-1.4	-1.8	-2.1	-0.1	-1.5	-1.2
7	-1.7	-1.3	-0.3	-1.7	-1.5	-1.8		-1.8	-1.6	-1.8	-1.8	-1.6	-0.5
8	-0.8	**1.0**	-0.6	-0.6	-1.6	-2.9	-1.7		-1.4	-1.9	-0.7	-0.2	-1.9
9	-0.1	-0.8	-1.3	**0.8**	-0.4	-2.0	-0.7	-1.8		-0.9	-1.3	-1.0	-0.6
10	-1.5	-1.2	-1.1	-1.9	-1.2	-1.0	**0.9**	-1.0	-1.4		-1.0	-0.9	-1.0
11	-0.7	-1.9	-1.6	-1.3	-1.9	**1.1**	-1.6	-1.0	-1.6	-0.6		-1.9	-0.3
12	-2.7	-1.5	-0.7	-2.8	-1.1	-1.9	-1.9	**1.1**	-2.0	-2.0	-2.3		-1.6
13	-1.2	0.1	-1.1	-0.3	-1.6	-0.9	-1.8	-2.0	**1.9**	-1.2	-0.8	-0.6	

Each row except row 7 has a unique entry that is substantially larger than the others in the row (larger by at least 0.6), and that is much closer to 1.15 than to -1.41. (In fact, these entries average 1.07, very close to the predicted 1.15.) If this entry occurs in row *i*, column *j*, it signals that in all liklihood in the plain text column *j* comes immediately after column *i*.

All the other entries are considerably smaller than any of the entries describe above, and average -1.33, close to the predicted -1.41. If one of these entries occurs in row *i*, column *j*, it strongly suggests that column *j* does not come immediately after column *i* in the plaintext.

Note row 7 has no entry indicating another column following column 7 in the plaintext. Presumably this indicates column 7 is the last column in the plaintext.

Likewise, column 5 in our matrix above has no entry indicating it comes after some other column. Presumably it comes first in the plaintext.

We can read off the presumed order of the columns, starting with 5:

5, 12, 8, 2, 13, 9, 4, 1, 11, 6, 3, 10, 7

Rearranging the ciphertext columns in this order, we get the plaintext.

5	12	8	2	13	9	4	1	11	6	3	10	7
t	a	k	i	n	g	a	t	i	p	f	r	o
m	a	l	b	e	r	t	e	i	n	s	t	e
i	n	s	c	i	e	n	t	i	s	t	s	h
a	v	e	f	o	u	n	d	t	h	e	s	m
a	l	l	e	s	t	m	o	s	t	e	a	r
t	h	l	i	k	e	p	l	a	n	e	t	y
e	t	u	s	i	n	g	s	t	a	r	s	t
o	b	o	o	s	t	t	h	e	p	o	w	e
r	o	f	t	h	e	i	r	t	e	l	e	s
c	o	p	e	s	t	h	e	n	e	w	p	l
a	n	e	t	i	s	a	b	a	l	l	o	f
r	o	c	k	c	o	a	t	e	d	i	n	f
r	o	z	e	n	o	c	e	a	n	s	r	o
u	g	h	l	y	f	i	v	e	t	i	m	e
s	t	h	e	s	i	z	e	o	f	e	a	r
t	h	i	t	i	s	m	o	r	e	t	h	a
n	t	h	o	u	s	a	n	d	l	i	g	h
t	y	e	a	r	s	a	w	a	y	i	m	p
o	s	s	i	b	l	y	o	u	t	o	f	r
e	a	c	h	f	o	r	f	o	r	e	s	e
e	a	b	l	e	h	u	m	a	n	t	r	a
v	e	l	b	u	t	t	h	e	m	a	n	n
e	r	o	f	i	t	s	d	e	t	e	c	t
i	o	n	s	u	g	g	e	s	t	s	t	h
e	u	n	i	v	e	r	s	e	c	o	u	l
d	b	e	f	u	l	l	o	f	m	i	d	s
i	z	e	d	t	e	r	r	e	s	t	r	i
a	l	p	l	a	n	e	t	s	n	o	t	t
o	o	d	i	f	f	e	r	e	n	t	f	r
o	m	o	u	r	o	w	n	w	e	t	h	i

What if we did not know in advance that the degree was 13.

The process of computing the matrix of average scores of column pairs is simple enough that (with a computer), we can perform it for many possible degrees.

For example, if we want to test degree 11, we arrange the ciphertext in 11 columns (omitting a partial column at the end), and compute the matrix:

	1	2	3	4	5	6	7	8	9	10	11
1		-0.6	-0.9	-0.8	-2.0	-1.5	-0.8	-0.8	-2.1	-1.9	-0.7
2	-1.0		-1.0	-1.0	-0.7	-1.0	-1.5	-1.6	-0.8	-0.9	-1.0
3	-1.2	-0.9		-1.1	-1.6	-1.2	-1.2	-1.2	-1.6	-0.4	-1.1
4	-1.3	-1.8	-1.3		-1.5	-0.5	-2.5	-0.6	-1.9	0.0	-0.7
5	-0.9	-1.2	-1.5	-1.7		-2.1	-1.0	-1.3	-1.0	-2.1	-0.3
6	-1.1	-0.2	-1.0	-1.5	-2.6		-1.1	-2.5	-1.0	-1.6	-0.8
7	-1.5	-2.3	0.1	-0.4	-1.7	-1.4		-1.1	-1.7	-1.3	-0.7
8	-1.6	-0.9	-1.1	-0.7	-1.2	-1.1	-1.6		-1.3	-2.5	-1.7
9	-1.0	-0.8	-2.0	-1.3	-0.8	-0.8	-0.8	-0.8		-0.9	-1.6
10	-1.6	-1.6	0.1	-1.3	-2.3	-0.4	-0.7	-0.9	-0.8		-1.0
11	-0.9	-1.1	-1.3	-0.9	-1.1	-0.9	-0.8	-0.8	-1.1	-1.7	

If we try to choose the largest element of each row, except one, we get the shaded entries in the matrix. (Even these choices cannot be correct, because in three cases, two elements lie in the same column.) In any case, the numbers are much too small, averaging -0.35 (considerably closer to -1.41 to 1.15).

So we would reject 11 as the degree of the transposition cipher.