

Part X

Proto-mechanical ciphers

Unit 120

Disk ciphers (cipher clocks)

Consider a device that looks like a pocket watch. Around the outer rim are symbols like letters and digits. Inside that is another ring of symbols. The outer ring has m symbols equally spaced around it, while the inner ring has n symbols, also equally spaced. The device has two hands like a clock. When the longer hand moves from one symbol on the outer ring to the next, the shorter hand moves from one symbol on the inner ring to the next. In other words, the hands are geared in the ratio $m:n$. The device can be used to encipher a text by rotating one of the hands clockwise until it points to the first plaintext symbol; the ciphertext symbol is the one to which the other hand points. The first hand is advanced to the next symbol in the plaintext; the second ciphertext symbol is indicated by the other hand. This continues until the full text is enciphered.

The *Wadsworth cipher disk* uses the outer ring for ciphertext symbols. It holds 33 symbols, the letters A-Z and digits 2-8. They can be removed and replaced in any order. The inner ring is for plaintext symbols, of which there are 26 (just the letters). Rather than having hands, in this device the two rings that rotate in the same direction under a single pointer. To encipher a text, the inner disk is rotated until the plaintext letter is under the pointer. The gearing rotates the outer ring to bring the ciphertext letter to the pointer. This operation is equivalent to the “cipher clock” described in the previous paragraph. The disks are always rotated in the same direction, presumably counterclockwise.

There is no evidence that there was ever more than one Wadsworth disk in existence. We do not know if it was used more than once. It was invented by Decius Wadsworth in America in 1817, more than fifty years before Charles Wheatstone reinvented it in England (too slow, Britain!).

The *Wheatstone Cryptograph* is better documented, especially by Wheatstone himself. Its outer ring holds the plaintext symbols, which are fixed in place. There are 27 of them, the space and 26 letters. The inner ring has 26 ciphertext symbols, which are the 26 letters in any order. The initial position before the first letter is enciphered is for the longer hand to point to the space and the shorter hand to point to the first letter of the mixed ciphertext alphabet. The hands are always rotated clockwise. To encipher a text, the hands are rotated clockwise until the longer hand points to the plaintext letter. The shorter hand points to the ciphertext symbol.



Wadsworth cipher disk. Photo from NSA.



Wheatstone Cryptograph. Photo from eBay.

In his paper describing his device (citation below), Wheatstone introduces a novel way of mixing the ciphertext alphabet from a keyword. It goes like this:

1. Write the keyword in a row.
2. Take the 26-letter alphabet and delete any letters that appear in the keyword.
3. Add the remaining letters in the alphabet under the keyword in rows with the keyword letters at the top of each column.
4. Duplicate letters in the keyword are deleted.
5. The mixed alphabet is read off in columns.

Here is an example using the keyword **WHEATSTONE**:

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

The mixed alphabet is

WBPHCQEDRAFUTGVSIXJYOKZNLM

The procedure for encipherment on the Wheatstone disk is

1. Remove all punctuation from the plaintext. Keep only letters and spaces.
2. Either:
 - a. Put an 'X' between any double letters in the plaintext. If the double letters are already 'XX,' then put a 'Q' between them.
 - b. Replace the second letter of doubles with 'X' or 'Q.' Be sure not to accidentally get a doubled 'X' or 'Q.'
3. If the plaintext does not end with space, then add one. The last ciphertext letter serves as a sort of checksum.
4. Position the longer hand of the device so that it points to the space on the outer ring of symbols and the shorter hand so that it points to the first letter of the mixed ciphertext alphabet.
5. For each plaintext symbol (letter or space):
 - a. Rotate the longer hand clockwise until it points to the plaintext symbol. Do not go all the way around. Let the shorter hand move on its own, according to the gearing.
 - b. Read the ciphertext symbol, which is indicated by the shorter hand.

Step 2 is necessary because there is no combination of ciphertext letters that the Wheatstone disk can decipher to a repeated letter, since the ciphertext alphabet is shorter than the plaintext alphabet. This is not true for the Wadsworth disk; for double letters on that disk, we would simply rotate 26 more steps for the second letter.

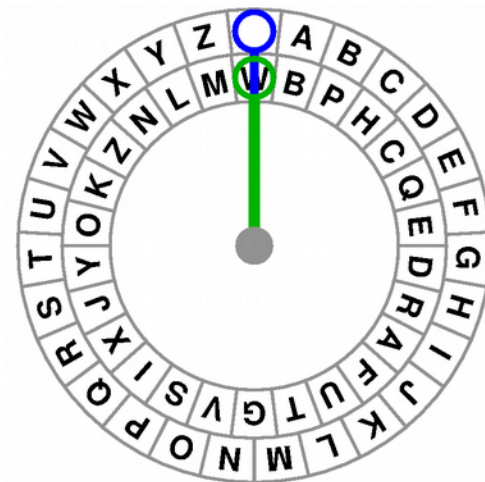
Let's look at a short example and encrypt this message with the Wheatstone disk. We will use the mixed alphabet from the example above.

SECRET MESSAGE

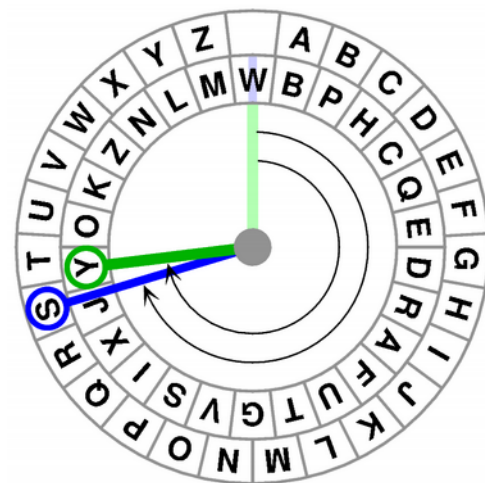
First, we have to prepare the plaintext:

SECRET_MESXSAGE_

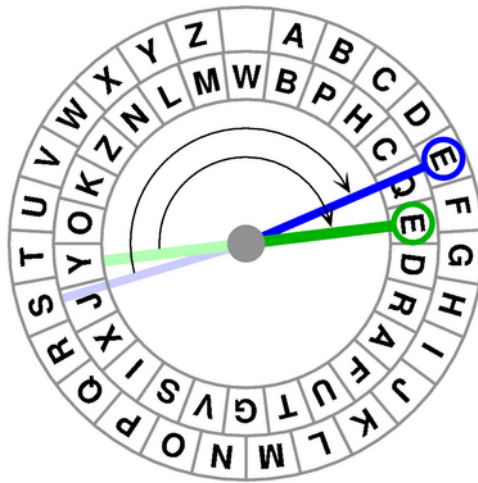
We begin the encipherment by putting the hands of the device in their initial positions.



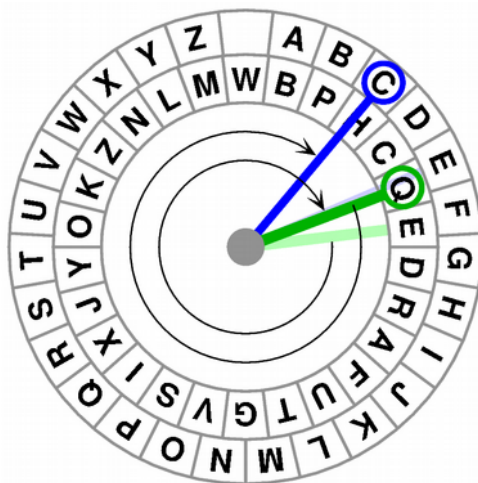
The first letter of the plaintext is 'S,' so we turn the large hand clockwise 19 steps until it points to 'S.' At the same time, because of the gearing of the device, the short hand also moves 19 steps and lands on 'Y,' which is the first letter of the ciphertext.



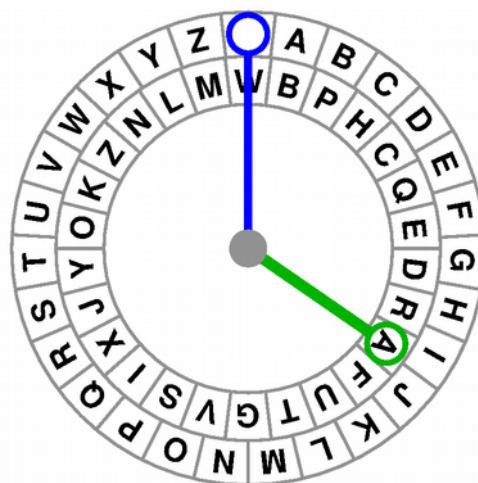
The second letter of the plaintext is 'E,' so we turn the large hand 13 steps clockwise to 'E,' while the short hand also moves 13 steps and lands on 'E.'



Since we always rotate clockwise, to encipher the next letter, 'C,' we have to turn the large hand most of the way around the circle. The short hand indicates that the next ciphertext letter is 'Q.'



The process continues, always clockwise, until the entire text is enciphered. By the time we get to the space at the end of the plaintext, the short hand has moved so far that it points to 'A' in the inner ring.



The full ciphertext is

YEQORNCXFLHMRVGA

The *Pletts Cipher Machine* was another cipherclock based on the Cryptograph. Its innovation was to mix both the plaintext and ciphertext alphabets, to make breaking its cipher more difficult.

The astute reader may have noticed that these three devices give us stream ciphers. Let's consider the Wheatstone disk. The internal state can be thought of as the positions of the two hands, $(h_{\text{long}}, h_{\text{short}})$. The initial state is $(0, 0)$. The cipher is factored into a stream cipher that corresponds to the disk using the key ABCDEFGHIJKLMNOPQRSTUVWXYZ, followed by a monoalphabetic substitution cipher M that uses the mixed alphabet as its key. For each character p_i from the plaintext, which we think of as a series of integers (space = 0, 'A' = 1, 'B' = 2, ...), the action of the encryptor is

$$\begin{aligned}x &= p_i - h_{\text{long}} \mod 27 \\h_{\text{short}} &= h_{\text{short}} + x \mod 26 \\h_{\text{long}} &= p_i \\c_i &= M(h_{\text{short}})\end{aligned}$$

For the Wadsworth disk, exchange the rolls of h_{long} and h_{short} , and replace 26 with 33 and 27 with 26. For the Wadsworth disk, we also need to specify that if x is ever zero, it should be changed to 26; this allows for the encipherment of repeated letters, which is never done with Wheatstone.

Let's now redo our example encipherment with the Wheatstone disk as a stream cipher. Start with an internal state of $(h_{\text{long}}=0, h_{\text{short}}=0)$. The first plaintext letter is 'S'=19, so x is $19 - 0 = 19$. The internal state changes to $(19, 19)$, and the output of the stream cipher is 19. Applying the monoalphabetic substitution gives us the first ciphertext letter, which is 'Y.'

W	B	P	H	C	Q	E	D	R	A	F	U	T	G	V	S	I	X	J	Y	O	K	Z	N	L	M
0	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

Next is 'E'=5, so $x = 5 - 19 = 13$ modulo 27. The internal state becomes $(5, 19 + 13 \mod 26) = (5, 6)$. The output is $M(6) = 'E.'$ Then comes 'C'=3. The internal state becomes $(3, 5)$. The next ciphertext letter is $M(5) = 'Q.'$ Etc.

A more interesting way to define the internal state S is as the total number of steps taken. From this point of view, the encryptor does these things for each character that it enciphers:

$$\begin{aligned}x &= p_i - S \mod 27 \\S &= S + x \\c_i &= M(S \mod 26)\end{aligned}$$

During World War II, the Danes used a derivative of the Wadsworth and Wheatstone devices, called *Urkryptographen* (Danish for “clock cryptograph”). Like the Wadsworth disk, it had two concentric rings of characters that rotated counterclockwise under a single fixed pointer, and the key was written on the outer ring. Like the Wheatstone Cryptograph, the plaintext alphabet contained a space which the ciphertext did not; therefore, double letters had to be disguised with ‘X.’ The letter ‘Q’ was used to indicate numbers, which appeared with some letters on the plaintext disk. The two letters ‘X’ and ‘Q’ are not used in native Danish words. Unlike either of the other two devices, Urkryptografen had three additional letters in each alphabet: ‘Æ,’ ‘Ø,’ and ‘Ü,’ so that the rings had 29 and 30 characters. The plaintext alphabet was also mixed, and there was a number of preprinted disks. One such plaintext alphabet was (where ‘_’ denotes the space)

_ØACUHDGVMKBÜIWRNPOZÆEXFSYJTLQ



Urkryptografen. Photo from Museum of Cypher Equipment in Fife, Scotland.

Reading and references

Charles Wheatstone, “Instructions for the Employment of Wheatstone’s Cryptograph,” *The Scientific Papers of Sir Charles Wheatstone*, The Physical Society of London, 1879, pages 342-347.
archive.org/details/scientificpaper00londgoog (the last two pages of the article were completely ruined by Google in that copy), books.google.to/books?id=CtGEAAAIAAJ

William F. Friedman, Several Machine Ciphers and Methods for their Solution, Riverbank Laboratories Department of Ciphers Publication No. 20, 1918, www.campx.ca/Several_Machine_Ciphers.pdf and www.marshallfoundation.org/library/methods-solution-ciphers

James Stanley, “The Wheatstone Cryptograph,” incoherency.co.uk/blog/stories/wheatstone-cryptograph.html

William F. Friedman, Six Lectures on Cryptology, www.nsa.gov/Portals/70/documents/news-features/declassified-documents/friedman-documents/publications/ACC15281/41785109082412.pdf

“Ciphers and Cipher-Writing,” Macmillan’s Magazine, XXIII, Feb 1871, pages 328-338, babel.hathitrust.org/cgi/pt?id=mdp.39015004979913;view=1up;seq=340

NSA file 41788379082740:
www.nsa.gov/Portals/70/documents/news-features/declassified-documents/friedman-documents/patent-equipment/FOLDER_515/41788379082740.pdf

Basic Cryptography, Dept. of the Army Technical Manual 32-220, April 1950, section 142 in chapter 11, www.nsa.gov/Portals/70/documents/news-features/declassified-documents/friedman-documents/publications/FOLDER_238/41748889078809.pdf

Niels Faurholt, “Urkryptografen (The Clock Cryptograph),” *Cryptologia* 27:3 (2003) 206-208, DOI: [10.1080/0161-110391891874](https://doi.org/10.1080/0161-110391891874); this article is available also at www.jproc.ca/crypto/crypto_watch.html

Greg Mellen, “Cryptanalyst’s Corner,” *Cryptologia* 8:1 (1984) 55-57, DOI: [10.1080.0161-118491858773](https://doi.org/10.1080.0161-118491858773)

David Kahn, *The Codebreakers: The Story of Secret Writing*, New York: Simon & Schuster, 1967, revised and updated 1996, pages 195-198.

Auguste Kerckhoffs, “La cryptographie militaire,” *Journal des sciences militaires* IX (1883) 5-39 and 161-191, www.petitcolas.net/kerckhoffs/crypto_militaire_1_b.pdf, www.petitcolas.net/kerckhoffs/crypto_militaire_2.pdf, section IV.

Donald W. Davies, “Charles Wheatstone’s Cryptograph and Pletts’ Cipher Machine,” *Cryptologia* 9:2 (1985) 155-160, DOI: [10.1080/0161-118591859870](https://doi.org/10.1080/0161-118591859870)

Programming Tasks

1. Write a function that mixes a ciphertext alphabet using the prescription of Wheatstone.
2. Implement an encryptor for the Wheatstone disk.
3. Implement a decryptor for the Wheatstone disk. Think about how to handle double letters in the ciphertext.
4. Implement a dictionary attack on the Wheatstone cipher.

5. Implement an encryptor for the Wadsworth disk.
6. Implement a decryptor for the Wadsworth disk. You do not have to worry about double letters in the ciphertext, since they cannot appear with this cipher. To handle double letters in the plaintext, advance the device 26 steps between them; the shorter hand will return to the same place, but the longer hand will point to a different ciphertext character.
7. Learn Danish and implement a simulation of Urkryptografen. Brug det for at skrive i sifer din favorite aventuren av Hans Christian Andersen. Skift din bakkeopstigningangreb på Wheatstone-kryptografen (Unit 123) at tilægge de tre nye bogstaverne, og brug angrebben din at løse siferteksten din.

Exercises

1. On average, how many characters are enciphered before the Wheatstone disk passes through its initial configuration? Use your monogram frequency table in your calculation.
2. Encipher this text with the Wheatstone cipher using keyword **CONVOLUTION**.

Viewing the Wheatstone Cryptograph as two rotating pointers is only one way of seeing this cipher. Another way is through the lens of modular arithmetic. The modulus for the plaintext is twenty-seven, while the modulus for the ciphertext is twenty-six.

3. This is the example ciphertext from Wheatstone's article. It was enciphered with the keyword **FRANCE**. Each paragraph is enciphered separately; i.e., the disk is reset to the starting position at the beginning of each paragraph. Decipher the message.

PZLSPQREQAJDITFBUFZOHQOSUQUODIKITORTWEZACMTPLERAUESKGSO
FGFDKHL SJIRKHFHMFADAYIVUOHAOBLNOGREJAIBKMPJZTMJABQCNFP
OMYHYRCZDCWBXUBZ

ZBILIJTEJYSPFDLCXETKQASOXOUNNODQJCWECLXPUIYIEMMCMSYVCFP
OKWCDEDVDAGLPEEKNAGVKMNUULSHXYXYVGFQPUYIORQKLPTCZHHK

ZBKUPVSWZWXAQXDREKTKQASOXOUIRSKOMFSTIIXGWTQJJVDYFNAHLS
IIXIAGQLZXVOGNHGRBUOHYZOOPWVYDDMQJKFMOBJPDYVRBAWKGWSJU
RJGITOWTVEZBHSOSLVUNBCHQSOTEIEBDQMGWHGJAMISXFIFBBPAVPE
SVCJUTAD

PZLPTYVXQXDTGLTTAF CVMHOMBINJKWVYAZOCQLAIUKFEGFNCNFIZHH
KVZYQUGLIVENKAHTRVFVEBWHWLRXCMLXWSYSYJHUF SFPOKEGZRXLUB
FT

4. Break this ciphertext with a dictionary attack. The keyword is a ten-letter English word. The cipher is Wheatstone's.

ZUKMNXQCDJKEADWCJVQHKCWNGGFVMAGQFE0COW0IXTVIAKJPNIJJXU
UQKCLBXVABOFIAATPQHIVKJQYUWZQRZKXWXYZVILALHDSSIGGMQQVK
ZVYCVZUPMHRTCTJXCXBIJGLNUVWEBTJZTOQLIKPRYEDCXIGDMET
YPOWFEBARSZZCOHWPABUIZCLWJJYQQKGPTTLWAIZUKXNGWDVHVDXUX
EXEJBWJPPSXPNPUHRHTILXRKSGXKWMCFZMBTVMNUDCDVLAFQRRM
GGFOUWAHVXHNHABCUPXKZOABAWMMZGFVNOQCKYIOVGXDBUARIYSTJG
YVGWLPLKEWXHJSZVIMSYFPL

5. Decipher this text with the Wadsworth disk cipher in which the ciphertext alphabet is not mixed.

TABLV3BTJWYCVKNRCOXFLSD2GTL6JUBUJNQ7CUXZG4AK4NT2LB07DM
XYLQDQVCK8OW3AL6T2GWOU3MCPR40DKVZK56APHABIZ36JUBUN2J8Q
5M7XM3B2MUC2K8YH4TAK6BMXM6DXGL24EF7XHLV7KBE2M3FMZMOERU
3CZG50SVDHZ46LAFPAYBVG4IOH4IL3S6WYADMXQ3CILMUKY0YQHULP
SAEWZ3I6CM6RBKXZN5P2DEHL56LBM0EJ6CHYGKPVZQBZ205RARJQG
U78K60BZDK6CKT4BMEUBYP38ZNFY27LYH26P3GKY7YG70CPBPSSDFT
6XYFHMNSCQR7V8BY5RW3L2LPAQ5F0UIU8H7K2LAFU8YRBZJYL5V5T
HVL5HITGXKRFRGJW

Unit 121

Attacking cipher clocks with cribs

If we know some part of the plaintext (a “crib”), we can use it to break the ciphertext. To make it easier, we will encrypt the crib with an unkeyed cipher clock before we compare it to the ciphertext, since the difference between a ciphertext and a text that was encrypted with an unkeyed device is a monoalphabetic substitution.

An example will make the technique more clear. This ciphertext was encrypted with a Wheatstone Cryptograph:

```
DGBRHDYDURNFTWLOMIBLSPDZCIJNSMBZXCSHDYCVMG0AZSAEGNVQLJRFTX
CWJHACZZMPEHMZNPZJYKMRXQG0UIUCRBOYQLOTOBUNZGJTEQB0FYLEHBOBZ
RRGW0HWJKCUKPPZCIQELLOUEJVAYVYPQWHE0BJNVGMAXRGLYYHNAHYCYFTGR
ILIJNSMHQFUWYYUEJVNFIIXNSDMXZSAHCNFVDVXVZKMCJWLXTBKJJGHHDRSK
CIMACCWTUKZVSRDGZSKCNUQQGCDKWKHVHIIFIXUMYIPWYSOPCGGSGHZBFWW
XPCEWEQMISKVXEMIZXSHAZEWOMRGQUTKCTKMPEUQJYLHTIDHFPERYUKSKON
NCEPLARCSBLMK0BCBSEFGXN
```

We have this short bit of its plaintext:

```
never wear out no matter how long it may stand in times
```

Before we use the crib we will encipher it with an unkeyed Cryptograph. The ciphertext alphabet for this process is just the unmixed alphabet. Before encipherment we need to disguise the double `tt` by replacing the second with `q`.

```
never wear out no matqer how long it may stand in times
```

Enciphered this is

```
OGXHUDAJGXGVBBIWJWLECRENVCKOADDXRALSFUSUNOWJAXGLYSIMFT
```

Notice that we did not add a space to the end before encipherment, as Wheatstone had recommended, since we do not know if the next character after the crib is a space.

Now we compare the enciphered crib to the ciphertext, one position at a time. Start at the beginning:

DGBRHD¹RYDURN²FTWLOMIBLSPDZCIJNS³MBZXC⁴SHDYCVMGOAZ⁵SAEGNVQLJRFTX . . .
OGXHUDAJGXGV⁶BBIWXJWLECRENVCKO⁷ADDXRAL⁸SFUSUNOWJAXGLYSIMFT

Notice that B in the crib matches up with F and T in the ciphertext. Also, A in the crib matches up with R, S, C, and Z. These things cannot happen with a monoalphabetic substitution, so we know that this is not the correct position.

The only acceptable line-up we can find is at position 107:

. . .TEQBOFYLEHBOBZRRGWOHWJKCUKPZCIQELLOUEJVAYVYPQWHEOBJNVGMAXRG . . .
OGXHUDAJGXGV⁶BBIWXJWLECRENVCKO⁷ADDXRAL⁸SFUSUNOWJAXGLYSIMFT

Now we can begin to reconstruct the key by tabulating the substitutions for the letters of the enciphered crib:

ABCDEFGHIJKLMN¹OPQRSTUVWXYZ
ERCLKABFGHIJMPQ??UVXYZWON?

The missing letters are D, S, and T. Missing three letters means that there are six possibilities. If the key is mixed in some less obvious way, then we would have to try all six. But in our example, we can guess immediately that the key is

WONDERCLKABFGHIJMPQSTUVXYZ

and the keyword is WONDERCLOCK. Usually, we also have to try all twenty-six ways of rotating the key before we find the correct key, since we do not know the position of the cipher clock's hands when the plaintext was encrypted at the point that the crib begins.

Deciphering with the reconstructed key gives us the plaintext, which you should recognize as a paragraph from *The Wonder Clock* by Howard and Katharine Pyle.

CLICK BUZQ WENT THE WHEQLS AND THEN TICK TOCK TICK TOCK FOR
THE WONDER CLOCK IS OF THAT KIND THAT IT WILQ NEVER WEAR
OUT NO MATQER HOW LONG IT MAY STAND IN TIMES GARQET DOWN I
SAT AND WATCHED IT FOR EVERY TIME IT STRUCK IT PLAYED A
PRETQY SONG AND WHEN THE SONG WAS ENDED CLICK CLICK OUT
STEPQED THE DROLQEST LITQLE PUPQET FIGURES AND WENT THROUGH
WITH A DANCE AND I SAW IT ALQ

Now let's work through a more complicated example. This ciphertext was also encrypted with a Cryptograph:

VDZYHDCDOFCLZSCPWTKEIVTTRNVGSYSCPUVEUJOWVFTZBLXDYXNZAU
NICZVYFHIKQFZLEMNLQIJHRCGOJSRODYLLCKDTWQUAOXTAVJ00YOGU
ZGBRZXZIRD¹XAGYBLHRGUINGCCNQHJSUWPLEUVD²DAWVORXIRJB³U00YW
FQOTXDFRDYUJYE⁴VXUAITGMLIIHKGAEFFVDAKXEHGSYXAVVAQCNECYK
ZOXWFXRRQIJYAZVCPZDUWSHHPAWLVFKDYNLRBTSCYREOBLPVVXZMCF
DUNQ⁵SXB⁶SRKAPMWCYHTBUERKAVCCARWYDNEHXS⁷WXBPHRACQIWIOYBX

PFMNUKGFFDNBXASYSXDHQFLDNWYTDGMCWLT PUXZFUIIALEKEILREZ
JHNP AEIBYHRNVGDLKUFECLXCGEWMEFVLCZBUONKUISJCSOFCLZTNES
DUWSXNLROKNGWCHBOUIZZATRKNNPQH XJMONSMZDWKGTBZKPDQFTIJ
NVXJRKZP

This crib is known to be part of the plaintext:

information exists of course but it is scattered

After we replace the double tt with tq and encipher with an unkeyed Cryptograph, the crib is

JOHQTPEXNTTGLEQABBJYQLOAGEFSOQJJQZKRAKSLWVOMBOCC

The first position that does not give a contradiction in regarding the monoalphabetic substitution is 177. This is actually the only position that works.

. . . YE VXUAITGMLIIHKGAEFFVDAKXEHGSYXAVVAQCNECYKZOXWFXRRQI . . .
JOHQTPEXNTTGLEQABBJYQLOAGEFSOQJJQZKRAKSLWVOMBOCC

We begin to reconstruct the key:

ABCDEFGHIJKLMN O PQRSTU VWXYZ
EFR?GSHU?VCKWLXTANYI?OZMDQ

There are three missing letters (B, J, and P) in the key, so there are six possibilities. Plus, there are twenty-six ways to rotate the key. That makes $26 \times 6 = 156$ total possibilities. But we don't need to try them all. Let's just try one and see what happens. Start with

EFRBGSHUJVCKWLXTANYIPOZMDQ

There are twenty-six ways to rotate this key, and we try them until we find something useful. Deciphering the ciphertext with the key as written above gives gobbly-gook:

IXUPCUFTPVDGPYCMDGBQIYDCPDVPQCPUDQSIPQCTPGUAHR SBVQTYR . . .

Next we try FRBGSHUJVCKWLXTANYIPOZMDQE and get more nonsense:

HWTBTESOUCFOXBLCFAPHXCB OCUOPBOTCPRHOPBSOFT GQRAUPSXQ . . .

We continue to rotate the key, and soon we arrive at TANYIPOZMDQEFRBGSHUJVCKWLX, which gives this plaintext:

THE NEQD FOR INXORMATION OF AN EOACT AND RELSBCMFADIB . . .

Aha! We are getting close. Most of the key is correct, but now we need to find the correct placement of B, J, and P. By swapping B and P we get the true plaintext, which is from *The Modern Clock*, by Ward L. Goodrich.

THE NEQD FOR INFORMATION OF AN EXACT AND RELIABLE CHARACTER IN REGARD TO THE HARD WORKED AND MUCH ABUSED CLOCK HAS WE PRESUME BEQN FELT BY EVERY ONE WHO ENTERED THE TRADE THIS INFORMATION EXISTS OF COURSE BUT IT IS SCATQERED THROUGH SUCH A WIDE RANGE OF PUBLICATIONS AND IS FOUND IN THEM IN SUCH A FRAGMENTARY FORM THAT BY THE TIME A WORKMAN IS SUFQICIENTLY ACQUAINTED WITH THE LITERATURE OF THE TRADE TO KNOW WHERE TO LOQK FOR SUCH INFORMATION HE NO LONGER FEQLS THE NECESQITY OF ACQUIRING IT

Notice that instead of trying 156 possibilities, with this method we only need to try at most $26 + 6 = 32$ possibilities.

Now that we have the plaintext, we can stop. But, no, we can't stop. Never stop solving puzzles. Let's go onward and get the keyword. The key we found was

TANYIBOZMDQEFRPGSHUJVCKWLX

It looks random, but notice the sequence Q, R, S, (T is missing), U, V, W, X, with one or two letters between them. Also see F, G, H, J, K, L. So the key was likely constructed with Wheatstone's prescription. We proceed by breaking the key into columns and getting the sequences each to appear on one row.

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

If we search our dictionary for TIMEP??C?, we find the only choice is TIMEPIECE.

Programming tasks

1. Implement a boolean function that detects whether two (short) texts could be related by a monoalphabetic substitution.
2. Implement the attack that uses a crib to break a ciphertext encrypted with the Wheatstone Cryptograph.

Exercises

1. Break this ciphertext that was encrypted with the Cryptograph. Use the crib "the average collector would be bored." What is the keyword?

XRIHAOZJMWGRABONYAJXBGPMWFMDFHGBDBYDPVHTBKATCDNUIXEENO
LSUHVCLRCFHRVUQFKBMLMSFGLHMAIAYSWAMNEOPTSJQJBPGITLRYLA
YCRJTGWPWAKWAZXSZKONBVNHLXFQDKSPUONMFEGQBQBCNXNUHXXTM

RVXPVUQFKCQLMNYXQUIKNGPHANTBVKIVDBAVWONAZGVEJAMSJKXLHQ
TDZFFPJMSFLKMXGYSKBTBIWTDNGXOCNUYOHEDACVHXYPKUGNDIJJ
GQRLSFQEOMMXHQSPRYUMBIFYCVFYGAQJWSJOISLFVESKJNIVMXQXEK
YMYXRVELATVXPNBOMEABVQNUUBVIVUYOZQXYETDKKGGWWEIPXNEXT
MUFXNESNZXUNJIHVAPMFKYZBKWSXIZVMSCDZEQVCBIGGWUAPMGWMZK
BIZPJURNLOKMUAQJGSNVKAODZMYJGJIMCNUCYKMVVLZCQFSLNWWDYA
KQHQQRVELDBYSGONJEREYNPKEXSAEUCRQJREGW0IOIUUKLADDEZQRO
NHZALWJPEYHMDKIYVYDPFZBSFNMDVEOSQTKDNYIAXUDYAKNQECUULN
LCJICNOOHMVBTCYQLSXSIBNESKTP0YCX0COLMKHXYFKCNWIPRDGMW
Q

Challenge

crib: NOT A SOUND BROKE THE SILENCE OF THE STILL NIGHT

DOW#GLGHUJ#VOWNQJLPNWBDZQDAQ\$R@CVZIXRMEFP#HIGNPAMGCIKGJTPCI
DQTCSLF\$UHAXNE#VNGJUK@WDITNJXNHYRKIYQGEQ\$NE@IUVHRT#\$KCLVKS
JE0XCLESNBKFMJAS\$LIOSCNVCLFKMWN@#GHPJOXEYVLOST#RHGOMFASYRT
F#AVPHFGUS@TQHUBIXYB#JLTCJRDUS#QND\$UEKBTCTAZPWLCXQYPSLUPG#J
@AMPXB@JIS\$VYJRTZAFRZUS\$MNOKASVKABVTBHK#ELPYF\$KXQSUJDOCHFLESG
CBZLDHXCZRZFENO#BC@YPGW@TGJVTANFVIYFKGY0IZSV@JCSLUPG#BUV#LNS
GHILJZTYNVZUFQHRKCESL\$ATUZE@IUOUKWCYKBHVOAUPF\$PSGSADHP@ZQ
M#TGJ#H\$N\$UQRHDFN\$JXE@IOQLMKRYJNWNQBSW@DROQT#VPENJ@YJJIU\$R@C
VHUXRIYAB0YQZGJDKZJ@QIDQTCH#PHXAHKQSTEU0#N@RSTYPBZTFWLZ\$CSZ
@CLVDP@RXE\$ZGKXJADKWITE@JMJPCPLYOAZXAQRBOWRIJSNDJ@YRXEFEHXO
HYR@Y\$ZCDKMRLAGJFTIEBNK#LIWNCXQYP#YJNWZKTBG0PTGBHPLI

Unit 122

Digram-counting attack on cipher clocks

This attack only applies if the length of the ciphertext alphabet is at least two more than the length of the plaintext alphabet. The reason that it works, if the text is long enough, is that when we encipher the next letter of a text, then the hands on the cipher clock move at most m steps (remember that m is the length of the plaintext alphabet, and n the length of the ciphertext alphabet). So if n is one greater than m , there are no double letters in the ciphertext, but all other digrams are possible. If n is two or more greater than m , then the missing digrams indicate what letter(s) is found on the key just before the last ciphertext letter that we wrote down.

To show how it works, we will work through three examples.

example 1: $n = m + 2$

Below is a ciphertext. It was encrypted with a cipher clock that uses this plaintext alphabet:

ABCDEFGHIJKLMNOPQRSTUVWXYZ

and this ciphertext alphabet, before it is mixed:

ABCDEFGHIJKLMNOPQRSTUVWXYZ◇◇

And here it is . . .

XAYIHVBS◇URMXE◇WBQPSTGXECXCXPA◇OCATLXVBEQSZYB◇DTAJSNEKGJDV
WEAXA◇GBNGQWJQXOJEGPXJLFXFESFMWDOWEKEJFPWE◇CABSYUWSKOLZDMH◇
BIBYTJQYUECVAUEOCSPKWLGM◇HIEWNEJKCUBSXMSEQVALMNTUBSNBMD◇OJE
QGADBPGYPHYQJVHBHIBTDKTAPCEUMI◇WEHNGZIDVQXZENWAQBPFHJKMSYOB
UORLGNBLZ◇QAWPRIEOVY◇RVYUFCHO◇UZJ◇DVUNAT◇NF◇UPXKOPYRSZ◇ST
JHKNSMPRDVXTGEKITAPLUEJ◇OVR◇SUPJPZKGNGDWCXIPJDIXCPBYOQMSLZU
JVNxED◇QG◇WENUGCN◇OFEPATLMPDWLCDEOATAP◇D◇Z◇CFORQLUZPXQ◇ZGOK
TA◇NKMPG◇MQLZUJVNB◇ADBNALQ◇E◇KLND◇XUF◇BL◇E◇GOBHDA◇DTAGXPXIY
UVBCYOJ◇TQD◇WDLIO◇NJKB◇TDRUZRNG◇H◇JW◇PSE◇NRD◇XUSLHZSRNXSDXZ
GKIJEUYXCPVOMHGHSLNULRZPU◇UNO◇EGEAPEXUFWEACWJYIEMLHJXNOWLIQ
YXKVXV◇GRCPYKJDP◇DJZGVPOVUBIQGAZIHUYX◇RLWBSL◇DSGTBEV◇PSKE
DLGRLORQYVJQWBGBZMDAWCUQ◇IKET◇PISD◇QPVSZGTLTDZODTUO◇MEKO◇FX

OMXCOGZIDVUNAXKRENOQ◇FHVAW◇UIVPSKGGJVRNGQVMDJYBPQBQ◇FHOMSFEC
WORYQSZLMK○ZFPSIJQCEHXSEHDHJQJGZOBPSBQMUGJOMSFECWORYQSZLMK○
ZFPSIJQCEHXUJ◇MHMPAEGBNXEIDUWCDK○OQYXKMSIVOPBZKQDRGWZSHBYTA
PLXB◇CSF◇SMAYMEJPTHJWGNBYN○UCZMBL○WCSMTGYJOUBHJQZF○ILIESECK
◇FXIVRXMSENJGQJQDVXECJSJU○UKSPAUFBOYVJVJMRVJ○BWCUKXEJPA◇ELV
JS○QR◇VSEASOGUIMCGDEUYXVYJYLNQEZQ◇PQRO○NKMZXOSVASNKLZJFAP
WCE○FJVGMKMXDZ◇RXMSNSMJNB◇XB○JEMNDI◇YSKWAP◇SRU◇H○JABHNQEOLE
DASYTCPYOSGMIVGOJQAKUCLXJRNUPAIHVBDU◇LMR○OTLGIZUF○KGXNRDLU
MIDZLCOBGZIJHGCYPEGTJ◇DUAVAVH○MJRJEYQDRGLHXT○BVF◇KSREJXCX◇R
V○RLGBQNBYPDNDKMXSLHUJHKHJBDGIUNTWBV○KTAJVGMDK◇AUYFZSNQGR
HBTXUJNKDT◇JRFEUMIHOBXAEILWTCZSJ○KTAPGIFI◇GD◇FS○VZLZGJBTCXB
○JEMNDUPKXEVL○E○MDWDGHIETAPGAZVR○SUGL○EYSWTBI○JZFAVVFV◇N◇R◇
FSFPJOWJ◇ILMIXVZOPBPS◇PA◇○KBSZ◇GPELOROXOHYNEHTJZ◇T○J◇TPKVZD
BPD◇OJHEJDHDGQEAWGNBJ◇OPCMKBR○IVSCOHYQJUGQ○XZKDGEEXGAZVR○SU
XAEYGZIUULT○XNXBPJNFUI○UTGLGKOYBLJWMKDFJWBVDRE◇FHVYBLSCSZSNK
◇XHQDEAGNBTJDVZUF○EUVYPKBSDJXAGAYSLUZG○VIAOHESYJT○HIW0XBCL
VQNR◇AXUXZNPBRQ◇XFHIEWDA◇QUDHIXYFZS○JYSPKBDHJSNOACE◇KECUIHV
YDNHVJDWXXMKDYPGEN◇FHJBEVCXUOELYJWMKDTETDSAUCUIECXF◇CXPHSZL
V○KVUXBJ○EKPDKHLH○PX○TGWTK○HBLXECTJO◇VZSVAWQJASZKXHSIJJZF○
EW◇UNKP○XQJUJEYLMNPSWIWBUSPFJUBDVZVFTXCFEGWP◇SNBMDWBC○QNPHO
BJRTBPUZJHCGQJBZ○DTQYXCKMXDGRMPYJRLRNG◇DJZVLICVHEJKPAVF◇EM○
VYQDUBDXUJWBYDUGYXFCOFEBLGJHUYQJHLT◇DTETPRI◇PJFC◇VF◇RVXAHUY
◇WQEUGWCJ◇EIAGBV◇AIDORQYQZLGMVLVQXYTXBQYNRDJ◇BZ○FZSY◇LZOTL
MDAPILIPYUI◇HBFSR○RVAV◇L○OMKVUZENUAQBPFDIVXDJSUYFOROSXMATL
GXPZ◇DSCLUXUIIMPKECUGTLRXSTYGZIF○PEWPIFYGOIJVAWDVL○UZPOWDLHE
XZT○EQLXCZDJUVCOPBHYMBZBZXGEUGPNDW◇HSE◇MXJRBSRNDGZH○JQCPE○
SUIIMNDKVEXUFCPTLDHMB◇V◇GW○UAPXNSZPGSAYPVHYXJWMLFZF○◇Y◇○CUZI
○ACVBSMAYQJVUGQMDXBPECUQ○NJVXVGBCP◇LIUWSZYXTQYUJ◇LXCVCWOUZT
○IZJBNYSBDIPSCZTXLEMNYQJVLIEANOKEBQYXFCXYVQWPDOALM○NGTJFPY
BLIPGCKMJDWEALPHOROVGHIEWDYJNDUZYOS○DLXMAJHJL○OMK○CDECD○K
TFPDXCOMPVHC◇YSMOKSZRXLTHJOE○UPKRUCGLUXBEAORXPHEQSHRXMKL◇IS
◇TDUG○PCDIVAI○WZ◇BMJ◇LSROUKL◇I◇LOXQTJUGCX◇SR◇CPYTLQDUBZDEVL
I◇FHFLFAEBSTPY○UK◇OWJLTDJZF○QP◇LUEHXVMCWJYQJNB◇WGBCTQGKUGL
PFXOASABRZ◇KMKDZTXSTAPYXJFXAEBNSGZIB◇IBFPSIPQUIJB◇HFCPNVEJ
◇YJHSEXBEQEQ○RVKXYVFSY○U○OQRMVWDQBWQUQPY○LNR◇FXE0JWJALFJGT◇
SL○TABIE○FNREJKVXPIV○RTXNOHJLSZJBDKVHIBTAIUQNUIPA◇ORQRZLDUB
THZOXSMJQYXCZPIEOJWJALSXHDHMG CJQZF○ME○BIFECUSYOTQXAVUZDTAG◇
HTBTCUKOQGDLFZC◇SMNEQSHFJMKL◇○XGWJUXSURVKX◇CEGMAUILWOSGAWOW
CEN0◇FHJKPYR◇EMZDVZ◇○NOVB◇HIBY○UY○PFNA◇○SY

The first step in the attack is to tabulate all of the digrams in the ciphertext. Remember to count *all* of them. The first four letters are XAYI; from them we get XA and AY and YI; don't forget that one in the middle. We don't need to count the digrams, just make a table of all the ones that exist.

AB			DB	EB	FB	GB	HB	IB	JB	KB		MB	NB	OB	PB	QB	RB	SB	TB	UB	VB	WB	XB	YB	ZB	◇B	○B
AC	BC			EC	FC	GC	HC	IC	JC	KC	LC	MC		OC	PC	QC	RC	SC	TC	UC	VC	WC	XC		ZC	◇C	○C
AD	BD	CD		ED		GD	HD	ID	JD	KD	LD	MD	ND	OD	PD	QD	RD	SD	TD	UD	VD	WD	XD	YD	ZD	◇D	○D
AE	BE	CE	DE		FE	GE	HE	IE	JE	KE	LE	ME	NE	OE	PE	QE	RE	SE	TE	UE	VE	WE	XE		ZE	◇E	○E

	BF	CF	DF				HF	IF	JF		LF		NF	OF	PF		RF	SF	TF	UF	VF		XF	YF	ZF	◊F	◊F
AG	BG	CG	DG	EG			HG		JG	KG	LG	MG	NG	OG	PG	QG	RG	SG	TG	UG	VG	WG	XG	YG	ZG	◊G	◊G
AH	BH	CH	DH	EH	FH	GH		IH	JH	KH	LH	MH	NH	OH	PH		RH	SH	TH		VH		XH	YH	ZH	◊H	◊H
AI	BI		DI	EI	FI	GI	HI			KI	LI	MI		OI	PI		RI	SI		UI	VI	WI	XI	YI	ZI	◊I	◊I
AJ	BJ	CJ	DJ	EJ	FJ	GJ	HJ	IJ			LJ	MJ	NJ	OJ	PJ	QJ	RJ	SJ	TJ	UJ	VJ	WJ	XJ	YJ	ZJ	◊J	◊J
AK		CK	DK	EK		GK	HK	IK	JK			MK	NK	OK	PK	QK		SK	TK	UK	VK		XK	YK	ZK	◊K	◊K
AL	BL	CL	DL	EL	FL	GL	HL	IL	JL	KL		ML		OL	PL	QL	RL	SL	TL	UL	VL	WL	XL	YL	ZL	◊L	◊L
	BM	CM	DM	EM	FM	GM	HM	IM	JM	KM	LM			OM		QM	RM	SM		UM	VM	WM	XM	YM	ZM	◊M	◊M
AN	BN	CN	DN	EN	FN	GN	HN		JN	KN	LN	MN			PN	QN	RN	SN		UN	VN	WN	XN	YN	ZN	◊N	◊N
AO	BO	CO	DO	EO	FO	GO	HO	IO	JO	KO	LO	MO	NO		PO		RO	SO		UO	VO	WO	XO	YO	ZO	◊O	◊O
AP	BP	CP	DP	EP	FP	GP	HP	IP	JP	KP	LP	MP	NP	OP		QP		SP	TP	UP	VP	WP	XP	YP	ZP	◊P	◊P
AQ	BQ		DQ	EQ		GQ	HQ	IQ	JQ	KQ	LQ	MQ	NQ	OQ	PQ		RQ		TQ	UQ	VQ	WQ	XQ	YQ		◊Q	◊Q
	BR		DR			GR	HR		JR	KR	LR	MR	NR	OR	PR	QR		SR		UR	VR			YR	ZR	◊R	◊R
AS	BS	CS	DS	ES	FS	GS	HS	IS	JS	KS	LS	MS	NS	OS	PS	QS	RS			US	VS	WS	XS	YS	ZS	◊S	◊S
AT	BT	CT	DT	ET	FT	GT	HT	IT	JT	KT	LT	MT	NT	OT	PT	QT	RT	ST		UT		WT	XT	YT	ZT	◊T	◊T
AU	BU	CU	DU	EU	FU	GU	HU	IU	JU	KU	LU		NU	OU	PU	QU	RU	SU	TU		VU		XU	YU	ZU	◊U	◊U
AV	BV	CV	DV	EV	FV	GV	HV	IV	JV	KV	LV	MV	NV	OV	PV	QV	RV	SV		UV			XV	YV	ZV	◊V	◊V
AW	BW	CW	DW	EW	FW	GW		IW	JW	KW	LW	MW	NW	OW	PW	QW		SW	TW	UW	VW					◊W	◊W
AX	BX	CX	DX	EX	FX	GX	HX	IX	JX	KX	LX	MX	NX	OX	PX	QX	RX	SX	TX	UX	VX	WX		YX		◊X	◊X
AY	BY	CY	DY	EY	FY	GY	HY	IY	JY		LY		NY	OY	PY	QY	RY	SY	TY	UY	VY		XY		ZY	◊Y	
AZ	BZ	CZ	DZ	EZ	FZ	GZ	HZ	IZ	JZ		LZ	MZ		OZ	PZ	QZ	RZ	SZ		UZ	VZ	WZ	XZ				◊Z
A◊	B◊	C◊	D◊	E◊	F◊	G◊		I◊	J◊	K◊	L◊	M◊	N◊	O◊	P◊	Q◊	R◊	S◊	T◊	U◊	V◊	W◊	X◊	Y◊	Z◊		◊◊
	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◊	◊◊	

What is important to us now is which digrams are *missing*. Of course, all of the diagonal entries are missing, since double letters cannot occur in the ciphertext for a device with this configuration. But look at the J column. The digram for JI is the only other one missing. Therefore, we know that I comes immediately before J in the key.

Here are all of the missing digrams, except for the doubles:

	BA						HA			KA	LA						RA							YA	ZA		
		CB									LB																
			DC											NC											YC		
					FD																						
																									YE		
AF				EF		GF				KF		MF				QF								WF			
					FG			IG																			
																QH				UH		WH		YH			
		CI								JI				NI		QI			TI								
											KJ																
	BK				FK						LK							RK					WK				
													NL														
AM													NM		PM					TM							

							IN					ON				TN								
													QO			TO								
														RP										
		CQ			FQ										SQ							ZQ		
AR		CR			ER	FR			IR							TR			WR	XR				
																TS								
																		VT						
											MU									WU				
																TV				WV				
							HW								RW						XW	YW	ZW	
																						ZX		
										KY		MY									WY			○Y
										KZ		NZ					TZ					YZ	◇Z	
A○																								

Now don't get confused, but to make things easier, we reverse each missing digram so that it is written in the way it might appear in the key. Here is the new table:

	AB					AH			AK	AL					AR						AY	AZ		
		BC								BL														
			CD								CN										CY			
					DF																			
																					EY			
FA				FE	FG			FK	FM				FQ					FW						
				GF		GI																		
													HQ			HU	HW		HY					
		IC					IJ			IN			IQ		IT									
								JK																
	KB			KF						KL				KR				KW						
											LN													
MA											MN	MP			MT									
						NI						NO			NT									
													OQ		OT									
														PR										
		QC			QF										QS						QZ			
RA		RC		RE	RF		RI								RT			RW	RX					
																ST								
																		TV						
										UM										UW				
																VT				VW				
						WH									WR						WX	WY	WZ	
																						XZ		
									YK	YM										YW				Y○

[illegible]

Now we look for any digram that appears alone in a row or column. Such a digram must be in the key. We find CD, FG, IJ, NO, MP, QS, HU, TV, Z \diamond , Y \circ , DF, EY, JK, LN, PR, ST, XZ, \diamond H, and \circ A. Once we decide to keep a digram in a row, we can eliminate the other digrams in its column; similarly, once we keep a digram in a column, we can eliminate the others in the same row. We can also eliminate the reversal of any digram that we keep. Here is what is left after we make those deletions:

[illegible]

Now we play the same game and look for digrams that are alone on a row or column in what remains in the table. We find GI, OQ, UM, VW, and WX. Eliminate those rows and columns to get an even smaller table:

LOOKUP ON THEM RATHER IN THE LIGHT OF OLD AND CONSTANT FRIENDS THAN AS MERE CHAIRS AND TABLES WHICH A LITTLE MONEY COULD REPLACE AT WILL CHIEF AND FIRST AMONG ALL THESE IS MY CLOCK MY OLD CHEERFUL COMPANIONABLE CLOCK HOW CAN I EVER CONVEY TO OTHERS AN IDEA OF THE COMFORT AND CONSOLATION THAT THIS OLD CLOCK HAS BEEN FOR YEARS TO ME IT IS ASSOCIATED WITH MY EARLIEST RECOLLECTIONS IT STOOD UPON THE STAIRCASE AT HOME I CALL IT HOME STILL MECHANICALLY NIGH SIXTY YEARS AGO I LIKE IT FOR THAT BUT IT IS NOT ON THAT ACCOUNT NOR BECAUSE IT IS A QUAIN OLD THING IN A HUGE OAKEN CASE CURIOUSLY AND RICHLY CARVED THAT I PRIZE IT AS I DO I INCLINE TO IT AS IF IT WERE ALIVE AND COULD UNDERSTAND AND GIVE ME BACK THE LOVE I BEAR IT AND WHAT OTHER THING THAT HAS NOT LIFE COULD CHEER ME AS IT DOES WHAT OTHER THING THAT HAS NOT LIFE I WILL NOT SAY HOW FEW THINGS THAT HAVE COULD HAVE PROVED THE SAME PATIENT TRUE UNTIRING FRIEND HOW OFTEN HAVE IS AT IN THE LONG WINTER EVENINGS FEELING SUCH SOCIETY IN ITS CRICKET VOICE THAT RAISING MY EYES FROM MY BOOK AND LOOKING GRATEFULLY TOWARDS IT THE FACE REDDENED BY THE GLOW OF THE SHINING FIRE HAS SEEMED TO RELAX FROM ITS STAID EXPRESSION AND TO REGARD ME KINDLY HOW OFTEN IN THE SUMMER TWILIGHT WHEN MY THOUGHTS HAVE WANDERED BACK TO A MELANCHOLY PAST HAVE ITS REGULAR WHISPERINGS RECALLED THEM TO THE CALM AND PEACEFUL PRESENT HOW OFTEN IN THE DEAD TRANQUILLITY OF NIGHT HAS ITS BELL BROKEN THE OPPRESSIVE SILENCE AND SEEMED TO GIVE ME ASSURANCE THAT THE OLD CLOCK WAS STILL A FAITHFUL WATCHER AT MY CHAMBER DOOR MY EASY CHAIR MY DESK MY ANCIENT FURNITURE MY VERY BOOKS I CAN SCARCELY BRING MYSELF TO LOVE EVEN THESE LAST LIKE MY OLD CLOCK IT STANDS IN A SNUG CORNER MIDWAY BETWEEN THE FIRESIDE AND A LOW ARCHED DOOR LEADING TO MY BEDROOM ITS FAME IS DIFFUSED SO EXTENSIVELY THROUGHOUT THE NEIGHBOURHOOD THAT I HAVE OFTEN THE SATISFACTION OF HEARING THE PUBLICAN OR THE BAKER AND SOMETIMES EVEN THE PARISH CLERK PETITIONING MY HOUSEKEEPER OF WHOM IS HALL HAVE MUCH TO SAY BY AND BY TO IN FORM HIM THE EXACT TIME BY MASTER HUMPHREYS CLOCK MY BARBER TO WHOM I HAVE REFERRED WOULD SOONER BELIEVE IT THAN THE SUN NOR ARE THESE ITS ONLY DISTINCTIONS IT HAS ACQUIRED I AM HAPPY TO SAY ANOTHER INSEPARABLY CONNECTING IT NOT ONLY WITH MY ENJOYMENTS AND REFLECTIONS BUT WITH THOSE OF OTHER MEN

You might recognize the text from *Master Humphrey's Clock* by Charles Dickens.

example 2: $n = m + 3$

For this example, we need to add another character to the ciphertext alphabet. This time, for each character in the key, we should find two missing digrams for the two characters that follow it. In which

order those two occur can be deduced from the choices of what characters might follow them. For example, if we have missing digrams AB, AC, BC, and BD, then we know that B and C are after A. Since C is after B, the order is ABCD.

Here is a ciphertext for us to break:

JEY@VHRIECWIBIRMBSVECN*LOTHSDIZ*DLPAF*LHUN@XCFXLACYUTHBPTCT
KVDENVWEXBVKMVGK#WG*GJVYSI@WAHOMY@YDZQ#JTERJBLHMRBQCXGUAMUB
HAPZG@AV#WSPDINPX#HEKEHVLSQLD#HVLV@WJZMUHYIMTCWH@CXKVDYOWUE
QZGRN#CA#QVECN*ISCVBL@FME#AWOA#TZVMPZR*CHXKICFSEHXZL@WELWK
VW#W@SYOQV*HPIPAFTXGJV*#MDHVSXLXLU@PHNKIJPBPXPZDJOTIXLRD#H
CEIZAOJEBGSOBOXFYPOMERWYSTFMKXRLYSP#@PBZQIB*B#GL*IJXGBVE*QY
RSAXYJCGQGWOWSWE0#YUAHVNFSA#HDPQFMRG*DHSKPGPIXGXLGERSCNF*SA
TJWNHEWXJPJFSI@CUR*VFKQXLQ*#IJUZBCWDHLXSMTJUXFSQRI#FUPM*BMA
XYPQOBUZNCLSMROPRW@WJYAWRGTLPG@RIYUCBYN#DXFIWUXOYKN#BOXCVOT
HUDOIN#DQSDS@AHZLVIUNG*GHYLBPGUEDUEYMANSDZFZRSQIE#Z@BOF*MDR#
BxBITBYJSDSHZ#PXUNTRLJOZBCQCLBCFLWZKW#IQEIAWRLUFR*BCBLENRKN
*NX@SMA#VWK@IJ*TEJBO@DTFODY@DYRAVYNHZF*NQYFIJY#FGQR#XQWEJPQ
CBJD*NGS@U#HMOADZWC@QTXGUTYZ#TWN*VRK*MEFCNQYHLYRGBN#*UBXJO
JEHKQVAFXQA#*W*K@PX#*BEMXHBRKGLBLFWSEWKRVEOZIBQS#MDV#IYS#D
@HWEHETUZR*@DX*XQBKNWKEADIJAY@WNGESGMOBCN#*BXFTCMBXLBIAORI
U@XZPLD*EQGJRLGBQZBPQU@DZCTCTKVIPRV#QZSJWBUX*VCP@V*FHBI*PCW
X@RJN@LOD@WJDTX@JZBPQW*BJSXAVNWSMEW#QZK#PKIPQAYN#MK*OIN#DLY
BKWRQIRXFKWJT*HJSDLWH#IBWNLKUMNBGIPXTYQT#MYFCMAHQTSMDG#OH*
@DPOXFKXKPBLJFSVHRCDDYD@GOBVKMCAIPED@RVENPVA#@AFEALZ*N*@WXU
TRSHQZKAT@QYJETBRSEXZUVGN@*GQRWF*TCDSFSCJAFZSIAILVGPOTZRJB
V#VN*#SCUAOFULUBIJQ@WOIUFWFSXKFJCJONZDRNYNKKYKJE#ALQSXJKSH
THURK*VLWSCNFOR#*MPT*@BCUVRHXYP*NOINCUDH#P#IJPRJBLMPDNW*NST
BUH#CJZV@RLB#QVCIPF*TKZSDNIESDSYZKNO@TP#LQG*OHMU#SNQ*JR@CZ
ENOIUIL*OP#PAFJEFXJMSM#YBDOFCLGZJUO*HZ*MXCLQGVX*BHXT@Q*JRK
*BRL@SCKZTGJOTUNIPXAERSILVSVWJRUYNRNLQGWUT@DJ#WCSKNLDLNDGNE
IAXYRVKGAIPQAHEKGEHPHY@LMVSIMUFGOXIUQ*#*UTMC*HLNL@BROKRZQJR
GJFRK*EDGKTS#MDPDNFRTYQVLVNFLLTB#CMXQDMGAZMNCNFOQJACAUKEYVY
IM#BVMFMKTFJACWP#MFXMAHTZ#J#QCSKGLAGMXTIUBLWOG#MHCMCPOS#MDW
IET#FIKEFK@XULTYGUTYKCQCSG@MZETUZRSLXKJQVOPBIPXKNSGRUVA#FIP
@JMRZT#SY*WX*OMU#SNWH@CXJZUGVPEIAPRQ#OZGLETLRHXEYEGHRNY*AHM
RZCFCILRJBVXMNCMLBJSDITRAMEMUYBFWZSEGC0UVP*#MCBRK#PH@JVACLM
Y@YAWYA*FRGUGKIPNYDMZTRUWCDYD#QZHKR#IRG@AJLAMYVPYRXINRCNOVD
WI*PZFXQZMFUXGVDKFULHZ@QJYABTFGNEGKRJPDRHVFIZDJNXAWRJZIDXOP
YASBQYKMLPMGNBRKVIBQWKIBVMPKSMCBFJPO#YF@LFTLYHMUN@QKJPVOEBP
IFSKNLDYH0IDWNOVOPGEZ*AOTV*J@RFXEHEMRVDR#J*FYG@#JPQJXZEGLIR
@#ZGLWRTBGB#S#WOSNPFVYCMUNRMNFL*KMPZDG*QK#QCXGNLDENOIUDZARH
JCUBE0ZJPJFYTZ@TYLHTAWQOFNOQUPLITEAYVM@GQE*NLDMOYKN@SYA#NC@
*EDQ#QOTHJNSNY@QFOUGMZ#P@WJDTX@JDJWKE*RKUCXLUZT*LPTUIQZGI*B
LJLO#OY@SXUIVSIU@AHYZTW*BQIBCZBDVSHTKBDQGSMDKH#DX@SYOQGESQVA
YOZCUP@DZNGRTRZAWJQL@JGWQDJC*B*GYCZEKVIFFANBILZAZBXRTUHMENX
G*DETZFN@AHPTWNFEADPDXMXNULHZH#ABMEJNCN@SYA#SKBY#IQ@E*VLFFVY
*RJAHEVMUGKNK@TPOPJXPXY#AMSHARLXRIMTPWRIQE*ANOQLKPHWOWKTLSC
SADZCMCXF@YZAZ#DXLI*GSHZDISNE@UALAON#P@BQWKIYAUNEBGYKRLSVFK

*@CVD*ALGWR@XPZATHZE#J#DLPEGSVHRCIDJV@ATOIR#FSPYKGEQYLTSCVB
 *DFMDCOQUBO*R@CL*AWCZBDJOYEJ*RFWLFVLNHQYEIACAQGABLXACVFNIA
 HRYGVCWOQLTYGYHKRC@Q@STFVO*LGNEWOALOTKJPZTBOISXUIK@VLEYGVBW
 PTFHRMWFLMGTB*DH#ZKW#JRQJQKBIV#JPHIWOWDJCVHPLIDLSISA*NEIUPZ
 DGRGTBP#*U@DZCGSMKVIR*ALVGMZTIQ@EN*SJYVKXISDWALWQAHJSRJMVS
 B#@ECUXHAVYNKDPEIAQIBX@WOI*CBPJMER#YVKFX@RJY#YPDHNKIUJODGTMK
 JWRDHUKXXLENEN@S#QZCJWBUIPKBCBWXSDJMUHXSQIDZIDJ#ADIULJFUT
 GJMZ@QJ@SMAN#CLHQF@GYTMXHI*LQ#CBORNB DYJACLMZUZNKX*BYMSMK#SV
 RGZL@ZEOKHSGINFEQZGRMZTI*TPV#UPYFHLQ#OPHZWRIUIRJBG#J#NF@UCB
 KUCKIYEOTGNZV#MDWDFXET@M*NTXY#I*FWEMV SIMUFHOW*VWLQTEADZPOHQ
 Z#JRCNCLFQRGSBYHCGWZBNTUXYCOQVUKFKTPQDTWY@M*MNWKTCCKZEHTSCUA
 JZADZQH#*STXEMENEHNSDVDKCNFHZARVEOZGLQTCBOZ@FVYAMFWATJFZ#*I
 J*J#W@BXZ#DJWS@QWDHCFLESUXFSFQLYOZI*HXK

Here is our table of digrams present in the ciphertext:

		CA		EA		GA	HA	IA	JA	KA	LA	MA		OA	PA	QA	RA	SA	TA	UA	VA	WA	XA	YA	ZA	*A	#A	@A	
AB		CB		EB		GB	HB	IB	JB	KB	LB	MB	NB	OB	PB	QB	RB	SB	TB	UB	VB	WB	XB	YB	ZB	*B	#B	@B	
AC	BC		DC	EC	FC	GC	HC	IC	JC	KC		MC	NC		PC	QC	RC	SC	TC	UC	VC	WC	XC	YC	ZC	*C	#C	@C	
AD	BD	CD		ED			HD	ID	JD	KD	LD	MD	ND	OD	PD	QD	RD	SD		UD	VD	WD	XD	YD	ZD	*D	#D	@D	
AE	BE	CE	DE		FE	GE	HE	IE	JE	KE	LE	ME	NE	OE	PE	QE		SE	TE	UE	VE	WE	XE	YE	ZE	*E	#E	@E	
AF	BF	CF	DF	EF				IF	JF	KF	LF	MF	NF	OF	PF	QF	RF	SF	TF	UF	VF	WF	XF	YF	ZF	*F		@F	
AG	BG	CG	DG	EG	FG				JG	KG	LG	MG	NG	OG	PG	QG	RG	SG	TG	UG	VG	WG	XG	YG	ZG	*G	#G	@G	
AH	BH	CH	DH	EH	FH	GH				KH	LH	MH	NH	OH	PH	QH	RH	SH	TH	UH	VH	WH	XH	YH	ZH	*H	#H	@H	
AI	BI	CI	DI	EI	FI	GI	HI			KI	LI		NI	OI	PI	QI	RI	SI	TI	UI	VI	WI	XI	YI	ZI	*I	#I	@I	
AJ	BJ	CJ	DJ	EJ	FJ	GJ	HJ	IJ		KJ	LJ			OJ	PJ	QJ	RJ	SJ	TJ	UJ		WJ	XJ	YJ	ZJ	*J	#J	@J	
		BK	CK	DK	EK	FK	GK	HK	IK	JK		LK	MK	NK	OK	PK	QK	RK	SK	TK	UK	VK	WK	XK	YK	ZK	*K		
AL	BL	CL	DL	EL	FL	GL	HL	IL	JL			ML	NL		PL	QL	RL	SL	TL	UL	VL	WL	XL	YL	ZL	*L	#L	@L	
AM	BM	CM	DM	EM	FM	GM	HM	IM	JM	KM	LM			OM	PM		RM	SM	TM	UM	VM		XM	YM	ZM	*M	#M	@M	
AN	BN	CN	DN	EN	FN	GN	HN	IN	JN	KN	LN	MN		ON	PN		RN	SN		UN	VN	WN	XN	YN	ZN	*N	#N		
AO	BO	CO	DO	EO	FO	GO	HO		JO		LO	MO	NO		PO	QO	RO	SO	TO	UO	VO	WO	XO	YO		*O	#O		
AP	BP	CP	DP		FP	GP	HP	IP	JP	KP	LP	MP	NP	OP				SP	TP	UP	VP	WP	XP	YP	ZP	*P	#P	@P	
AQ	BQ	CQ	DQ	EQ	FQ	GQ	HQ	IQ	JQ	KQ	LQ		NQ	OQ	PQ		RQ	SQ		UQ		WQ	XQ	YQ	ZQ	*Q	#Q	@Q	
AR	BR		DR	ER	FR	GR	HR	IR	JR	KR	LR	MR	NR	OR	PR	QR		SR	TR	UR	VR	WR	XR	YR	ZR	*R		@R	
AS	BS	CS	DS	ES	FS	GS	HS	IS	JS	KS	LS	MS	NS	OS		QS	RS		TS		VS	WS	XS	YS	ZS	*S	#S	@S	
AT	BT	CT	DT	ET	FT	GT	HT	IT	JT	KT	LT	MT	NT	OT	PT	QT	RT	ST		UT			XT	YT	ZT	*T	#T	@T	
AU	BU	CU	DU		FU	GU	HU	IU	JU	KU	LU	MU	NU	OU		QU	RU	SU	TU		VU	WU	XU	YU	ZU	*U	#U	@U	
AV	BV	CV	DV		FV	GV	HV	IV	JV	KV	LV	MV	NV	OV	PV	QV	RV	SV	TV	UV				YV	ZV	*V	#V	@V	
AW	BW	CW	DW	EW	FW	GW	HW	IW	JW	KW	LW	MW	NW	OW	PW	QW	RW	SW	TW	UW	VW				ZW	*W	#W	@W	
AX	BX	CX	DX	EX	FX	GX	HX	IX	JX	KX	LX	MX	NX	OX	PX	QX	RX	SX	TX	UX	VX	WX				*X	#X	@X	
AY	BY	CY	DY	EY	FY	GY	HY	IY	JY	KY	LY	MY	NY	OY	PY	QY	RY	SY	TY	UY	VY	WY	XY				*Y	@Y	
AZ	BZ	CZ	DZ	EZ	FZ	GZ	HZ	IZ	JZ	KZ	LZ	MZ	NZ	OZ	PZ	QZ	RZ		TZ	UZ		WZ	XZ	YZ			*Z	@Z	
A*	B*	C*	D*	E*	F*	G*	H*	I*	J*	K*	L*	M*	N*	O*	P*	Q*	R*		T*		V*	W*	X*	Y*	Z*		*#	@*	
A#	B#		D#	E#		G#	H#	I#	J#	K#		M#	N#	O#	P#	Q#	R#	S#	T#	U#	V#	W#	X#	Y#	Z#		*#	@#	
		C@	D@	E@	F@	G@	H@	I@	J@	K@	L@	M@	N@	O@	P@	Q@	R@	S@	T@	U@	V@	W@	X@	Y@	Z@	*@	#@		

[illegible]

We can **eliminate** all non-listed digrams from rows and columns that contain two of the digrams in our list, since we know that any other digrams in such a row or column is accidentally rather than necessarily missing. One such column contains **GI**, **HI**, and **OI**; since **GI** and **HI** are in our list, we know that **OI** cannot be. The table becomes

AN END EVERYTHING WAS SO DEATHLIKE THE MAN WHO CAN LIVE IN THE SAME HOUSE WITH ONE OF THESE CLOCKS AND NOT ENDANGER HIS CHANCE OF HEAVEN ABOUT ONCE A MONTH BY STANDING UP AND TELLING IT WHAT HE THINKS OF IT IS EITHER A DANGEROUS RIVAL TO THAT OLD ESTABLISHED FIRM JOB OR ELSE HE DOES NOT KNOW ENOUGH BAD LANGUAGE TO MAKE IT WORTH HIS WHILE TO START SAYING ANYTHING AT ALL THE GREAT DREAM OF ITS LIFE IS TO LURE YOU ON INTO TRYING TO CATCH A TRAIN BY IT FOR WEEKS AND WEEKS IT WILL KEEP THE MOST PERFECT TIME IF THERE WERE ANY DIFFERENCE IN TIME BETWEEN THAT CLOCK AND THE SUN YOU WOULD BE CONVINCED IT WAS THE SUN NOT THE CLOCK THAT WANTED SEEING TO YOU FEEL THAT IF THAT CLOCK HAPPENED TO GET A QUARTER OF A SECOND FAST OR THE EIGHTH OF AN INSTANT SLOW IT WOULD BREAK ITS HEART AND DIE IT IS IN THIS SPIRIT OF CHILDLIKE FAITH IN ITS INTEGRITY THAT ONE MORNING YOU GATHER YOUR FAMILY AROUND YOU IN THE PASSAGE KISS YOUR CHILDREN AND AFTERWARD WIPE YOUR JAMMY MOUTH POKE YOUR FINGER IN THE BABYS EYE PROMISE NOT TO FORGET TO ORDER THE COALS WAVE AT LAST FOND ADIEU WITH THE UMBRELLA AND DEPART FOR THE RAILWAY STATION I NEVER HAVE BEEN QUITE ABLE TO DECIDE MYSELF WHICH IS THE MORE IRRITATING TO RUN TWO MILES AT THE TOP OF YOUR SPEED AND THEN TO FIND WHEN YOU REACH THE STATION THAT YOU ARE THREE QUARTERS OF AN HOUR TOO EARLY OR TO STROLL ALONG LEISURELY THE WHOLE WAY AND DAWDLE ABOUT OUTSIDE THE BOOKING OFFICE TALKING TO SOME LOCAL IDIOT AND THEN TO SWAGGER CARELESSLY ON TO THE PLATFORM JUST IN TIME TO SEE THE TRAIN GO OUT AS FOR THE OTHER CLASS OF CLOCKS THE COMMON OR ALWAYS WRONG CLOCKS THEY ARE HARMLESS ENOUGH YOU WIND THEM UP AT THE PROPER INTERVALS AND ONCE OR TWICE A WEEK YOU PUT THEM RIGHT AND REGULATE THEM AS YOU CALL IT AND YOU MIGHT JUST AS WELL TRY TO REGULATE A LONDON TOMCAT BUT YOU DO ALL THIS NOT FROM ANY SELFISH MOTIVES BUT FROM A SENSE OF DUTY TO THE CLOCK ITSELF YOU WANT TO FEEL THAT WHATEVER MAY HAPPEN YOU HAVE DONE THE RIGHT THING BY IT AND THAT NO BLAME CAN ATTACH TO YOU SO FAR AS LOOKING TO IT FOR ANY RETURN IS CONCERNED THAT YOU NEVER DREAM OF DOING AND CONSEQUENTLY YOU ARE NOT DISAPPOINTED YOU ASK WHAT THE TIME IS AND THE GIRL REPLIES WELL THE CLOCK IN THE DININGROOM SAYS A QUARTER PAST TWO BUT YOU ARE NOT DECEIVED BY THIS YOU KNOW THAT AS A MATTER OF FACT IT MUST BE SOMEWHERE BETWEEN NINE AND TEN IN THE EVENING AND REMEMBERING THAT YOU NOTICED AS A CURIOUS CIRCUMSTANCE THAT THE CLOCK WAS ONLY FORTY MINUTES PAST FOUR HOURS AGO YOU MILDLY ADMIRE ITS ENERGIES AND RESOURCES AND WONDER HOW IT DOES IT I MYSELF POSSESS A CLOCK THAT FOR COMPLICATED UNCONVENTIONALITY AND LIGHTHEARTED INDEPENDENCE COULD I SHOULD THINK GIVE POINTS TO ANYTHING YET DISCOVERED IN THE CHRONOMETRICAL LINE AS A MERE TIMEPIECE IT LEAVES MUCH TO BE DESIRED BUT CONSIDERED AS A SELFACTING CONUNDRUM

IT IS FULL OF INTEREST AND VARIETY I HEARD OF A MAN ONCE WHO HAD A CLOCK THAT HE USED TO SAY WAS OF NO GOOD TO ANY ONE EXCEPT HIMSELF BECAUSE HE WAS THE ONLY MAN WHO UNDERSTOOD IT HE SAID IT WAS AN EXCELLENT CLOCK AND ONE THAT YOU COULD THOROUGHLY DEPEND UPON BUT YOU WANTED TO KNOW IT TO HAVE STUDIED ITS SYSTEM AN OUTSIDER MIGHT BE EASILY MISLED BY IT FOR INSTANCE HE WOULD SAY WHEN IT STRIKES FIFTEEN AND THE HANDS POINT TO TWENTY MINUTES PAST ELEVEN I KNOW IT IS A QUARTER TO EIGHT HIS ACQUAINTANCESHIP WITH THAT CLOCK MUST CERTAINLY HAVE GIVEN HIM AN ADVANTAGE OVER THE CURSORY OBSERVER BUT THE GREAT CHARM ABOUT MY CLOCK IS ITS RELIABLE UNCERTAINTY IT WORKS ON NO METHOD WHATEVER

The text is from *Clocks* by Jerome K. Jerome.

example 3: both alphabets mixed

For this example, we use a cipher clock with the twenty-six-letter modern-English plaintext alphabet and this twenty-eight-letter old-English ciphertext alphabet (before mixing):

abcdefghijklmnopqrstuvwxyzðþ

To make things easier, we also specify that mixing is done like this:

keywordðþabcfghijklmnpqstuvxz

Both alphabets will be mixed, so after the digram-counting attack, we need to break a monoalphabetic substitution cipher.

Here is the ciphertext:

ienbzuvucvxpðykyozkpujwajbpbllqpdgjhtxðwitskourixmjtggwdnah
tptkpevznyðatzjkfwkxzpaðvtnudhaopdbpcjuxdmðpkefigyoctdnppaz
bifwxhvpððynrdekzdzðeuhðrxðpinojwzðefwmbmvdujknizpðcucðhlm
nlbmdxixjwhdxkpampceufbqlcigvjuxzlhqjhticlacxubrcienpðdbuzmj
sgarsoenupzvsvrtpqwjnðaizbnvtxqjosotarctauðinzikefgbpxnvtxrj
puctðvtðadedofkezhpvpukonmirafnðfvstmiozbiezospjdukwnywzg
pcleuvrlyeopnxpmsqdsuxvkmszpeghtspclnrxydxpybzomqoelqclehðs
pqbynqiwmðdwglsgvgpðpjkboxuadwnjebzmjrpvwpslfilojueueðctlðz
evzhuyqvzbgvoepajwkðrfqbknðziompvpfvtoqhdpmqarcðrnðdprexlmw
psuypphqjekiyzpakabdeqdzbiczdtypsyudumnrjmfppbudnjhcmngxiuv
tvðpzuowlpkcarhiyzmngxczbsjsomodpbkwpjvrnðnlðbzuvculðyko
iozklpmpcsarnhpweibwlbiswglbrfoepwzlbgbrcmnkiedwnxfsysjreb
dvchlðptzpydbwctvðyrdrkpaxpghsnjhunzðpvnuvfvtstjupdkxfwðjrve
tnlzbizmutlqdwqdfgetaragxhðwkryzlyiuðvðmrfpvkyfperpmdqreiw
xzgmsxiutðedqvxoirsfjpaukbpiahagxiuvtxstcsgkonmirarcjrvetppa
nsxienrzwyrknfwpkwjtgðeplrhmkhfgtxrðugvðkcedncmteufdupegpx

[illegible][illegible][illegible]

[illegible]

Now we add **dn**, **su**, **vw**, and **pa** and eliminate the conflicts.

[illegible]

We can add jk , and eliminate je and jh .

[illegible]

				ef					em										
							fj							ft					
					gh				gm										
													hs	ht					
		id																	
							jk												
								kl											
						lj		lm	lo									lð	
						mi													
	nc				ng	nh													
						oh				op									
	pc										pq								
											qr								
									rp			rs	rt						
														su					
tb																		tz	tð
															uv				
																vw			
																	wx		
										xo								xy	
	yc						yj			yo	yq							yz	
				zf					zm										
				ðf						ðo									ðp
pa																			

Now we pick up ce and eliminate cg and cy.

	ab							aj			ao								
		bc														by		bp	
			ce																
										dn									
				ef					em										
						fj								ft					
					gh				gm										
													hs	ht					
		id																	
							jk												
								kl											
						lj		lm	lo									lð	
							mi												
	nc				ng	nh													
						oh				op									
	pc										pq								
												qr							

Looks like we're stuck. Let's see what we can do with this list of key digrams:

ce dn id jk kl mi ng qr su uv vw wx pa

We can get these key fragments:

ce jkl midng qr suvwx pa

Remember the form of the mixed alphabet. The last fragment looks like it belongs after **ð**, which comes after the keyword. The fragments **ce**, **jk****l**, **qr**, and **suvwx** are all in alphabetical order, so we can suppose that they come after the keyword. Fragment **mīdn****g** is definitely not in alphabetical order, so it must be part of the keyword. So far, we have

...midng...pa...ce...jkl...qr...suvwx...

Since **t** is missing from **SUVWX**, we can suppose also that **t** belongs in the keyword, but we do not know if it comes before or after **midng**. Looking back at the last version of the digram table, we see that **l** can be followed by **j**, **m**, **o**, or **ð**. We already know it cannot be **j**. Since **m** is in the keyword, it should not be **m**. And since **ð** comes between the keyword and the rest, it cannot be **ð**. So we can, with good reason, decide to put **o** after **l**. Furthermore, the table contains only **op** and **rp** in one column. So **p** is in the ordered part of the key, since **o** and **r** are. The table also has **pc** and **pq**. So let's put **p** between **o** and **q**. And in the ordered part of the key, nothing should come between **r** and **s**.

...midng...õpa...ce...jklpqrsvwx...

The table contains only `x0` and `xy` on one row, and since `0` is already after `1`, `y` must come after `x`. But we can't yet place `z` safely at the end.

...midng...đpa...ce...jklpqrsvwxy...

Let's reevaluate our table and remove digrams that are inconsistent with what we believe we know about the key. Does this remind you of the game Clue™?

[illegible]

RQFGJPDBABMFUAFRKIORIEODEFLFGCDQLHFKRCFBMFFXDABFQIFRKBMFIOR
IEDBDAABDOOQFIFAAGCYBRORREKRCBMFDQJFSFQJFQBDQVFQBDQRARKBMFW
FDLMBJCDVFGQJRKBMPFIMGQDIGOFAIGSFPFQBBMFKDCABRKBMFAFPGYAFF
PIRPSGCGBDVFOYBCDVDGOGYRQFKGPDODGCWDBMBMFCGDADQLRKMFGVYORG
JAHYPFGQARKCRSFAGQJSU00FYIRU0JAUCFOYCFIRLQDZFBMFSRAADHDODBY
RKUADQLAUIMGQGCCGQLFPFQBDQCFVFCAFGAGARUCIFRKABFGJYSRWFCQFVF
CBMFOFAABMFUAFRKBMDAJFVDIFDAQRBFCFIRCJFJHFKRCFDBAGAARIDGBDRQ
WDBMMYJCGUODIGQJSFCSFBUGOPRBDQRPGIMDQFADQBMFPGQUAICDSBARKCD
WQIGGQJDBAUAFDQGIORIEUADQLAUIMGSFCSFBUGOPRBDQRWMFFOPFCIUCYK
DOOFJGAGIORIEFAIGSFPFQBDQBMFGABCRQRPDIGOIRJDIFARKGOKRQARBMF
WDAFEDQLRKIGABDOFIGBMFAFIRQJDQVFQBDQRBMGBRKBMFPFIMGQDIGOFAI
GSFPFQBMGASCFQBFJRQFRKBMFPRABBGQBGODZDQLRKSCRHOFPAWDBMRUB
JRUHBBMFICRWQGGQJKRODRBBYSFRKFAIGSFPFQBGSSFGCABRHFBMFKDCABIR
PSODIGBFJPFIMGQDIGODQVFQBDRQEQRWQBRBMFFUCRSFGQPDJJOFGLFADBM
FCGOJARUCWMROFGLFRKPGIMDQFPGEDQLYFBQRBCGIFMGAHFFQKRUQJFDBMF
CRKGABFGJYFVROUBDRQRKAUIMFAIGSFPFQBARCRKBMFDCDQVFQBDQRDQFUC
RSFBMRULMBMFGABCRQRPDIGOIORIESRWFCFJHYGWGBFCWMFFOGQJLRVFCQF
JHYGQFAIGSFPFQBODEFJFVDIFMGJHFFQFOGHRGCBFJDQIMDQGKRCAFVFCGO
IFQBUCDFAHFKRCFBMFKDCABGSSFGCGQIFRKRUCIORIEAWFPUABQRWCFMFGC
AFGCFVDAFJABRCYRKBMFRCDLDQRKBMFIORIEGADBMGAHFFQAULLFABFJHYC
FIFQBCFAFGCIMFARQBMFMDABRCYRKLFGCDQLGQJRQIMDQFAFGQJRBMFCGAB
CRQRPDIGOPGIMDQFAGKBFCBMDAWFAMGOOKRCBMFKDCABBDPFSCFAFQBFVDJ
FQIFBRAMRWBMGBBMDAABRCYDAIUCDRUAOYCFQGBFJBRBMGBRKBMFSFCSFBU
UPPRHDOFRQFRKBMFLCFGBIMDPFCGARKAIDFQIFBMGBIGPFKCRPDBAPFJDFV
GORCDLDQBRSGYQDPSRCBGQBSGCBQPRCFQFQFQBFJFVFORSPFQARKFQFC
LFBDIAGQJBMFKRUQJGBDRQARKBMFCPRJYQGPDIADBDAGIUCDRUAPDXBUCFG
OOBMFPRCFARHFIGUAFBGQLOFJDQFXBCDIGHOYDQDBWFAMGOOKDQJBMFPRAB
DPSRCBGQBGQJFGCODFABCFFKCFQIFABRBMFUAFRKBMFPGLQFBDIIRPSGAAD
QBMFWFABDBAFFPABMGBDQCFVDADQLBMFMDABRCDFARKIORIEWRCEGQJBMFP
GLQFBDIIRPSGAABMFAFIRQADJFCGBDRQARKSFCSFBUGOPRBDQRQJFVDIFAPG
YSCRVDJFARFPFPUIMQFFJFJFVDJFQIF

We have to break this text as a monoalphabetic substitution. When we do, the plaintext is

THE HISTORIES OF THE MECHANICAL CLOCK AND THE MAGNETIC
COMPASS MUST BE ACCOUNTED AMONGST THE MOST TORTURED OF ALL
OUR EFFORTS TO UNDERSTAND THE ORIGINS OF MANS IMPORTANT
INVENTIONS IGNORANCE HAS TOO OFTEN BEEN REPLACED BY
CONJECTURE AND CONJECTURE BY MISQUOTATION AND THE FALSE
AUTHORITY OF COMMON KNOWLEDGE ENGENDERED BY THE REPETITION
OF LEGENDARY HISTORIES FROM ONE GENERATION OF TEXTBOOKS TO
THE NEXT IN WHAT FOLLOWS I CAN ONLY HOPE THAT THE ADDING OF
A STRONG NEW TRAIL AND THE ERADICATION OF SEVERAL FALSE AND
WEAKER ONES WILL LEAD US NEARER TO A BALANCED AND
INTEGRATED UNDERSTANDING OF MEDIEVAL INVENTION AND THE
INTERCULTURAL TRANSMISSION OF IDEAS FOR THE MECHANICAL
CLOCK PERHAPS THE GREATEST HINDRANCE HAS BEEN ITS TREATMENT
WITHIN A SELFCONTAINED HISTORY OF TIME MEASUREMENT IN WHICH
SUNDIALS WATER CLOCKS AND SIMILAR DEVICES ASSUME THE

NATURAL ROLE OF ANCESTORS TO THE WEIGHT DRIVEN ESCAPEMENT
CLOCK IN THE EARLY TH CENTURY THIS VIEW MUST PRESUME THAT A
GENERALLY SOPHISTICATED KNOWLEDGE OF GEARING ANTEDATES THE
INVENTION OF THE CLOCK AND EXTENDS BACK TO THE CLASSICAL
PERIOD OF HERO AND VITRUVIUS AND SUCH AUTHORS WELL KNOWN
FOR THEIR MECHANICAL INGENUITIES FURTHERMORE EVEN IF ONE
ADMITS THE USE OF CLOCKLIKE GEARING BEFORE THE EXISTENCE OF
THE CLOCK IT IS STILL NECESSARY TO LOOK FOR THE INDEPENDENT
INVENTIONS OF THE WEIGHT DRIVE AND OF THE MECHANICAL
ESCAPEMENT THE FIRST OF THESE MAY SEEM COMPARATIVELY
TRIVIAL ANYONE FAMILIAR WITH THE RAISING OF HEAVY LOADS BY
MEANS OF ROPES AND PULLEY COULD SURELY RECOGNIZE THE
POSSIBILITY OF USING SUCH AN ARRANGEMENT IN REVERSE AS A
SOURCE OF STEADY POWER NEVERTHELESS THE USE OF THIS DEVICE
IS NOT RECORDED BEFORE ITS ASSOCIATION WITH HYDRAULIC AND
PERPETUAL MOTION MACHINES IN THE MANUSCRIPTS OF RIWN CA AND
ITS USE IN A CLOCK USING SUCH A PERPETUAL MOTION WHEEL
MERCURY FILLED AS A CLOCK ESCAPEMENT IN THE ASTRONOMICAL
CODICES OF ALFONSO THE WISE KING OF CASTILE CA THE SECOND
INVENTION THAT OF THE MECHANICAL ESCAPEMENT HAS PRESENTED
ONE OF THE MOST TANTALIZING OF PROBLEMS WITHOUT DOUBT THE
CROWN AND FOLIOT TYPE OF ESCAPEMENT APPEARS TO BE THE FIRST
COMPLICATED MECHANICAL INVENTION KNOWN TO THE EUROPEAN
MIDDLE AGES IT HERALDS OUR WHOLE AGE OF MACHINEMAKING YET
NO TRACE HAS BEEN FOUND EITHER OF A STEADY EVOLUTION OF
SUCH ESCAPEMENTS OR OF THEIR INVENTION IN EUROPE THOUGH THE
ASTRONOMICAL CLOCK POWERED BY A WATER WHEEL AND GOVERNED BY
AN ESCAPEMENTLIKE DEVICE HAD BEEN ELABORATED IN CHINA FOR
SEVERAL CENTURIES BEFORE THE FIRST APPEARANCE OF OUR CLOCKS
WE MUST NOW REHEARSE A REVISED STORY OF THE ORIGIN OF THE
CLOCK AS IT HAS BEEN SUGGESTED BY RECENT RESEARCHES ON THE
HISTORY OF GEARING AND ON CHINESE AND OTHER ASTRONOMICAL
MACHINES AFTER THIS WE SHALL FOR THE FIRST TIME PRESENT
EVIDENCE TO SHOW THAT THIS STORY IS CURIOUSLY RELATED TO
THAT OF THE PERPETUUM MOBILE ONE OF THE GREAT CHIMERAS OF
SCIENCE THAT CAME FROM ITS MEDIEVAL ORIGIN TO PLAY AN
IMPORTANT PART IN MORE RECENT DEVELOPMENTS OF ENERGETICS
AND THE FOUNDATIONS OF THERMODYNAMICS IT IS A CURIOUS
MIXTURE ALL THE MORE SO BECAUSE TANGLED INEXTRICABLY IN IT
WE SHALL FIND THE MOST IMPORTANT AND EARLIEST REFERENCES TO
THE USE OF THE MAGNETIC COMPASS IN THE WEST IT SEEMS THAT
IN REVISING THE HISTORIES OF CLOCKWORK AND THE MAGNETIC
COMPASS THESE CONSIDERATIONS OF PERPETUAL MOTION DEVICES
MAY PROVIDE SOME MUCH NEEDED EVIDENCE

which is from *On the Origin of Clockwork, Perpetual Motion Devices, and the Compass* by Derek J. de Solla Price. The key for this substitution is

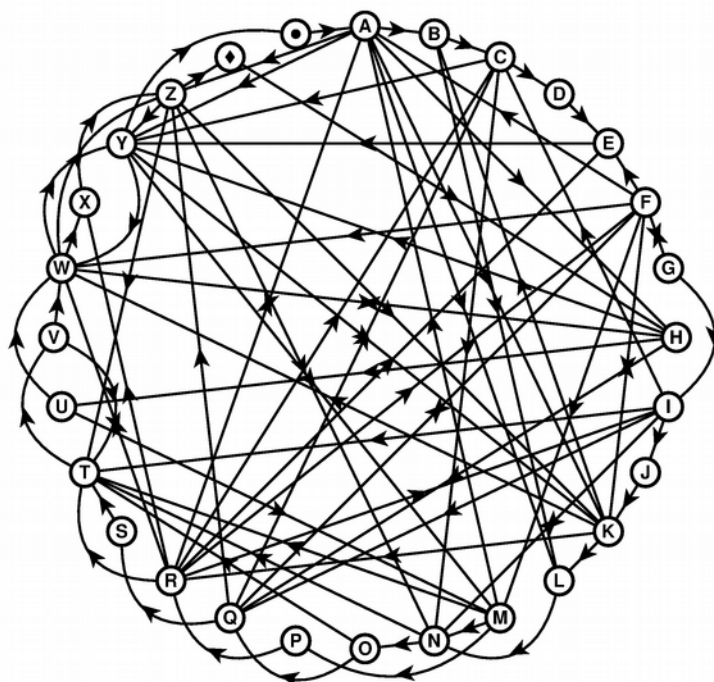
GHIJFKLMDNEOPQRSTCABUVWXYZ

If we invert this key, we obtain the mixed plaintext alphabet:

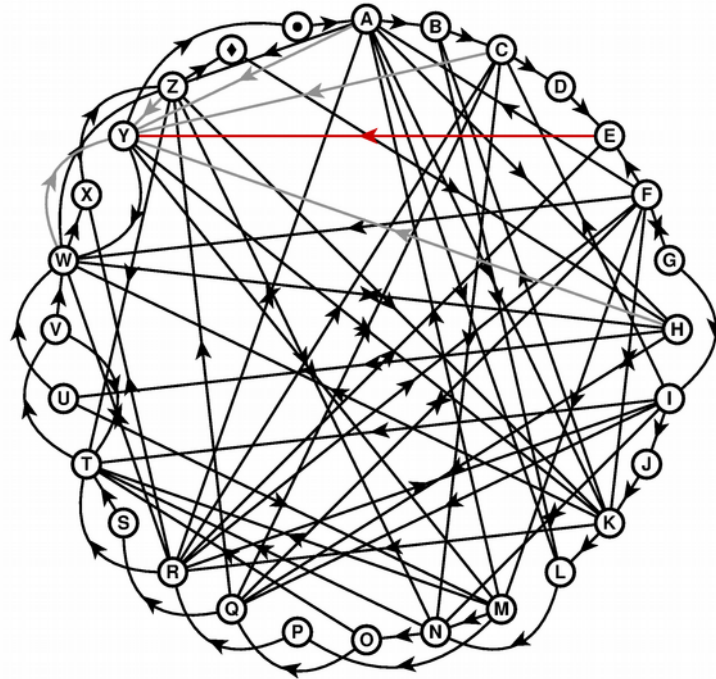
STRIKEABCDFGHJLMNOPQUVWXYZ

When working with pen and paper, it may be easier, especially when $n=m+2$, to do this attack with a directed graph. A *directed graph* is a graph whose edges are directed (have a direction). A *graph* is a set of vertices and a set of edges. A *vertex* is a point, but it can be drawn anywhere; we don't care about coordinates in this kind of graph, and vertices can be slid around on the page. An *edge* is a line from one vertex to another. In a directed graph, each of those lines has an arrow on it to show its direction. Think of a map in which cities are connected by one-way roads.

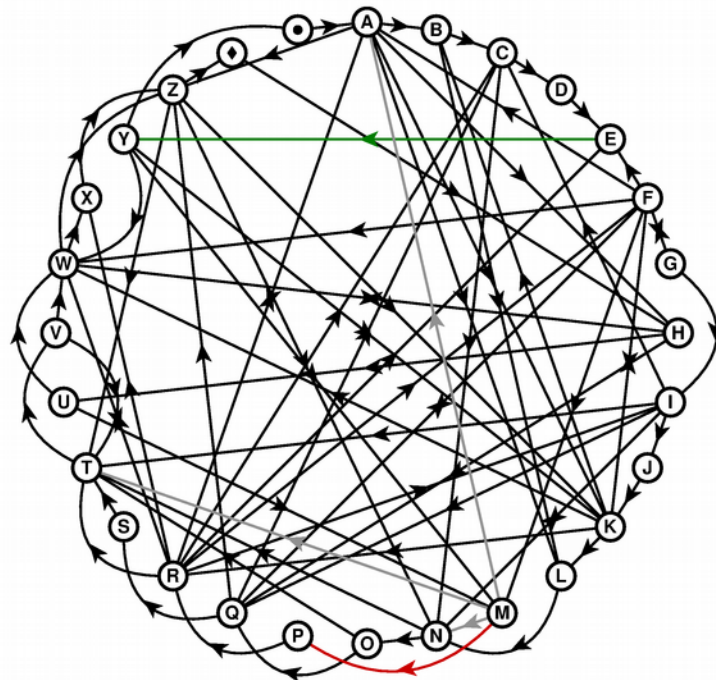
Let's redo the first example with a directed graph. If we take the reversed missing digrams and use each to define a directed edge, we have this graph:



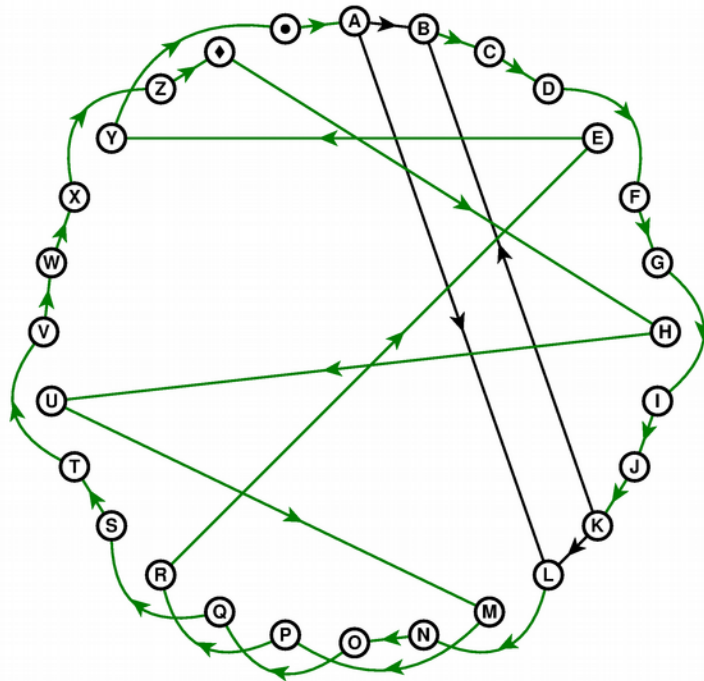
It is quite a mess. But notice that E has only one line directed outward from it. That line ends on Y. Since E must come immediately before Y in the key, we know that no other letter comes immediately before Y, so all other lines that end on Y can be removed.



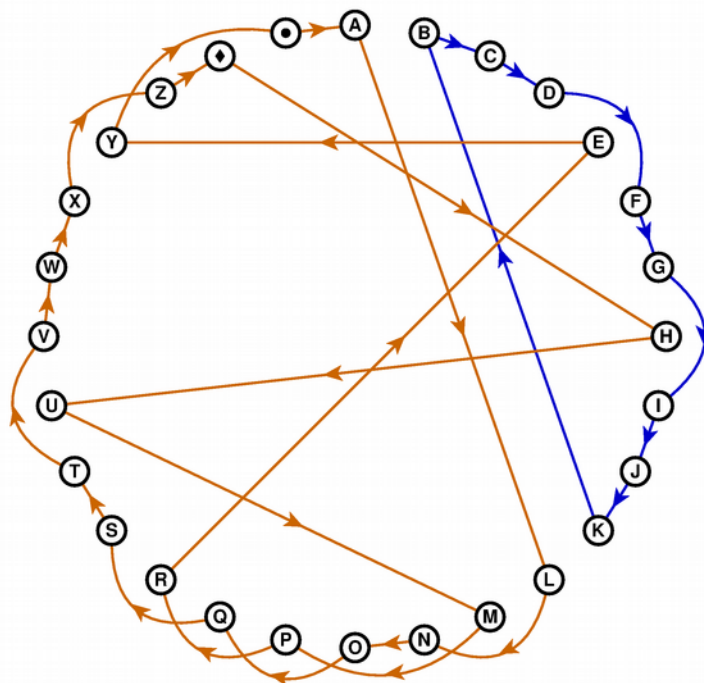
Also notice that P has only one edge that ends on it, from M. Therefore all other lines that start on M can be removed.



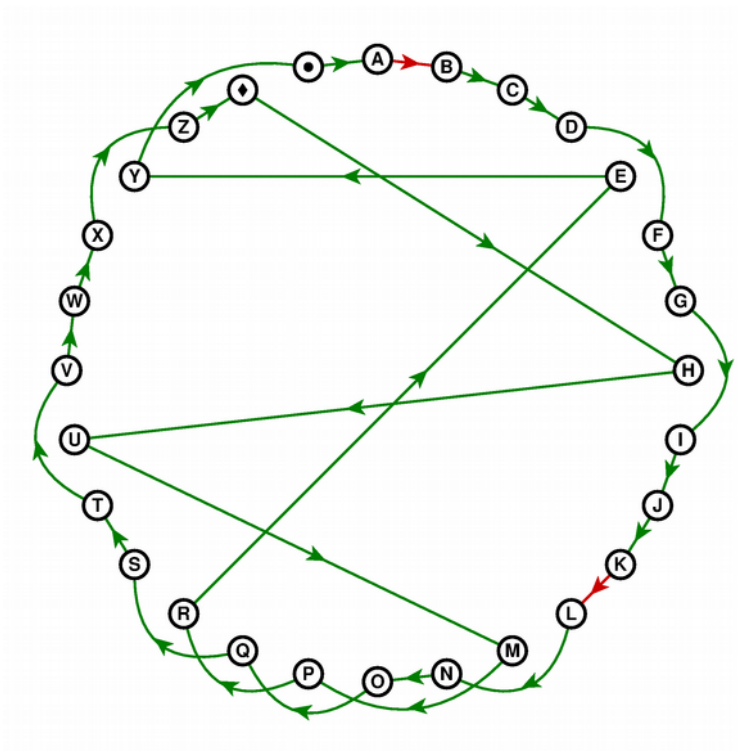
After a lot of this sort of work, we eventually get to the following graph, in which we can make no more eliminations. Did I say “pen and paper”? I meant “pencil.” The kind with an eraser.



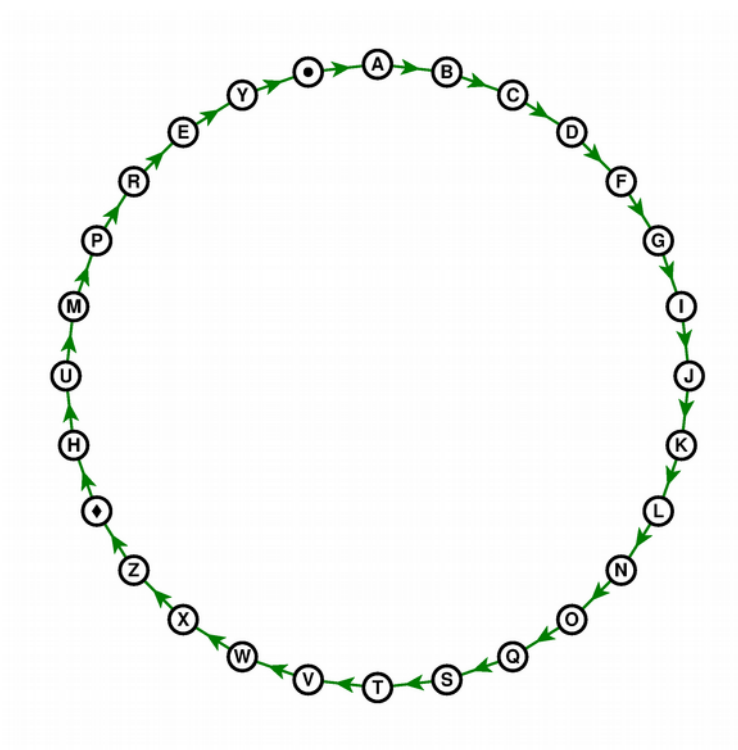
At this point, we have a choice to make. Either we keep A to B and K to L, or A to L and K to B. If we choose A to L and K to B, we see that the graph has two separate circuits (loops):



We can't have two loops, since we want one key to fit on the outer ring of the cipher clock. With the other choice, we get that one loop:



If we untangle the graph (remember that coordinates of vertices don't matter, and so vertices can be slid around), we see the key just as it would be set into the ring of the device:



Programming tasks

1. Implement the attack in the case $n=m+2$. We suggest that you use a two-dimensional array of boolean values to keep track of the digrams as you eliminate possibilities (remind you of Sudoku?). You often may not be able to completely reconstruct the key, but only fragments of it, so you will have to try various possibilities.
2. If you are so inclined, extend your attack for $n=m+1+k$, with $k \geq 1$. (It will become obvious why we write it that way.)

Exercises

1. You know what to do.

BT2QLFQDZSUYQBKQROWL1YFEZFOU3SEDPYFIA1XADSLHWZACIDGLVZ
DFMEJT3PXZPH1TC1GTIV2FOQHEYAQVD3JWYM2UX3KEZPGVQC13QVBM
VPGUKMLZLFJBLG1MQROSIW1BVNCMEUXLDXC02BYDBCRESVZUJMLHZF
DACID2XGAXH2KIRLXJGSVOET3NP1ROWFHZIG2TWLPOULTMPYRNATJ2
AYJE2CXBSFLSV0DFTENRZP01SWR3YKTF1MG2IAHUNPYI1KBH2DJVA1
RACWSWU2NCIRB1GB1BVOWPL2GZLYP1S2RUMPQG2MCXKGPUFTXPOUAJ
FYMNUFR2BPYMLHJQMYAKZBCZJKBJ2RLYJUMTMQSVGYQRWFLC3KSWEJ
RUXRZCMD1QW2T2XWUEM1HMTF3BF3FYLSAKPUKCE1UYLGVBQVGORJT
CUSWEJSWLM13SEKV2G2W2AYDNQFUQVIL2QTHOZ1C2UM2VDTAR0FXC2
JCWAQAIAREY2GCZLHT2DRWLROCHMFYATFEVPKTDNX23JVNLFSXMB
JLJSYCTPUHWBUPH12Y1CAORSWTJMEHI1AC3B3FLIROPQKSUNDZMCVY
1STBQTBW2QFIPBPROHUEZRNCU3AI12YQAWHOUJ1UXMA3JMZLYCERWZ
LYP1SUYVOMUL3YBKPTCWX3SDCKVJXC3J1QWNI12DUROYNRGISM
X23JVNLFSXMDUPBC2KGPUEZFD3LILBLJNTLXSYM3UIABVCLYCE2DIP
FDB3NVRKMZNGOZCUGPEILZD3IOSKIVNZSGTJKRCFYDABUG1BZTKT3B
ECNUIDQFPDP1HVRXJEDYWPOM13WJ1QW2TLKZNUH2NCJMKQITIRBK
MLZ3AF1EIUVDXQ1AMSTOR2MNBVNCMERTXANPI1CZBPXETYZMBUEXEC
HI1JAY1GB1IEGKTDNJ3KBH1SCZPDIKHKOBYIEZFD3NJHDV12UWCROL
GUJAUWRAY1MTGITJES2RLYAUM3CZMB30VRXJEBZLHVMNC3F0ZQ2TF
1UJQVI3WM3AXTLDBPJUKTIZMC1TNQXWDMYDWPXMGA1EZFD3JFGTV2G
2W2AYJENMB3MZBOQJ2RL3WHUACEDVDMRN2WLKWHKPVSWVBKMOY2CDB
DC2BZO3HI1JA3JBOLGUVKIMOWQZFKUKTLY10VIJ2MBC3G2IANSLRTB
QYC3TWDWXJSB3M3QFTGBOWPL2F1CUXIAYT1P2IRA2J2PTCSZVNUJ1M
KQOMZRSWPQEGJATUIRAPGCLXZPF2RFXUYULGMYAQVQMF0QR2KEGVZEF
TXZNW3DKQR12LHYMNUE2BPYMRHQEM3A1FEGHGL2GHZP01SCZWBEJ2
MBEHDXYWRUNJPSL1PLBZ2TWSGXSHJQYVKA0FYBOAFTUVKI30WZ2NOF
J3UYGLB1STLV2RBKMCWDRMTUQTKQLCGTDYPYOEKY2QSFGRLQWECFPE
S2GRZEMX23INCTUOVMD3FWXFHZ12DSGNYG2QYLMVOHTF1UHJVIEMA3
ULQGNSUPNXWJUNFSYRGIMD3FKJZSWEBL3T3KINYMNZKBRACRTZBXLU
2YKACK1LYCERWZ1J3UJTA1WBEIYTI3GVFWG01R12BGVDZSWZNUH2Y
IDYSYWVBMH2FKUH2VMTJOLSBWP2IWUDGLQ0F2KVIDRBYSHOWF3KAX1
D1HGKEQUGHI3PWVLSDWYKDI30FOQHEY1QNPS3KI1DCZOLBXFRTCHWP
XOX3SCLQ3JLYCE23TNJYN1SBLTRMBYIEZFD3QOZGVHJER2WPZGLHI3
1B1FT1LERPXNH2FIREFAUCEF0SILSJGHDXZFN03RD3KI2UWRVQHJXL

HWRIDHMYZ1LSV0VAWZM3YGLUMNSFOQH1CWRUNCT3JFKVXYHDNCGSHN
ZNB3MUR0QD2RL3U1TACZMNLILSJG2JRANWDUYLOFYMRATCHUSUDLUE
MAYGZXWBES2IOXJKYQGP3RFNOC2GVKXCWLQ3NACXLI1CKIOYKSWPV0
EJETMA23GYZLEVGKSI20YHUGZXMDUHOBKMCW1BPFOYROFYF1POTHWRI
DFEVYCUKMRKGRZKF3IUDGHSWR3TIJFMTMICXAQLD3JWYCGZROZNAEQ
CSYA3JBOWRP3PNCPXCLTVYQWUILSJGUCLUDTU3RS2MYAPZNTWXADJR
KQOCNPFLQSYFQJ2IOXJKYQVPLDUFTZBYUCWG013H2AKVXGRTCXWAZP
1QCJSJHBVGPWE23HKXYIUH23BQDNDJLT1XGP1EZFD3JFGTVBZB2LTJ
ETFNAQ2DSBWUXR3KCNXZSUVDQSES2TPLFEQRCMEJYWERXB3EAX03VX
GYRFZDXJYMABTRLGPSEIKTGFXNWIUIOC1NPDRZBFKJPNSDZAE3NY1P
L12FZSNPGX2QCLPQBJLIXODNBH22R1G3AUVHSESGTLAJL1KRH1EGNW
RWTLQUDF0VP3XWPSUEGMX3HOWF3KAYGZX0VAWZB3HNX3ROMDUVWFTN
BIYUCWE3NBKUDF0VXTFXNUDQKNQCPSEICLEYOTBEDRZ1CND1FSTVAW
GEIJFNORIFWL2CT3VDQWVBKILBDJMEHWRIDFTAX0ZABNDZEMXDVNXQ
VHPRDZBFPC3KGAUBW1DTU3KINYWCIOSDWBQ3JK1CWI1LDHMYQ0UXPJ
FJPIUFEQRGMD3FIPBM2HCWBUPH12BGVMKTGGGXK1CFYHDCMZ12VLNJ
NWZP01SCVKTEP3KPBHAUEZFDM3JFGTVDAMZR2WPGVQC1TGTB0QD2FN
JWZNTLQROSIVWFJCUSCGN03SMZBJ0ITKECFGCKS3EHI1JAQ2D1HGKP
RGF1JUX0BHGOEY1RIV2WLKWVHU2EFTXWPKRTMLBKTJ1L3SHPFD2DK
HIPGKWHMNBQZB3VFIAMDUPTKB3NBFPKJYHRMT3RHA2TIXGJPVZNSH
DS3I2DGYBOUVRWR3L3ZSHVDAYPHXUEGACZTYVNUGOYNCZJCRLVJETM
A3FRBXEQRE2IPZP01SDKPRTIGSJ2HEDSWHPOAL1HNVI3XPDGMWHUOC
PZ3UFAOL2KTDNXXBXYTAV2JWNBYUCWRG3SEKVPKXBAPRUZNUHU1CNP
FLQFLOCNXXBHUEYA1SYJ2FOQHETZFDST3KQ0YPHZ1HNPGKHZMB2KYBO
VAWZKEMPYAY1SEXMDUPB3KAZEDYWHXNOQM0QH22W2HTUEVP3SIEMKA3
JBOMERGKEYAUVDXQAXRFYCPXPFNHQA3MHTKIVOGXBXQDSLHWRIDGVT
WTEJT2FIREFAUVDXQNP1RWFT0B3VER1DFNDKRDKIVIDRLRDSWGQCFI
FKYEYGJ2

Unit 123

Hill-climbing attack on cipher clocks

There is a hill-climbing attack that we can perform on a ciphertext that was encrypted with the Wheatstone Cryptograph. It relies of the fact that we can factor the Wheatstone cipher into one with an unmixed ciphertext alphabet followed by a monoalphabetic substitution. We therefore start with the attack on the monoalphabetic substitution from Unit 28. To avoid becoming trapped in a local maximum, we need to use a margin of error that allows downward steps some of the time. We must build a new tetragram-frequency table, which will be built from a large text enciphered with the Wheatstone using an unmixed alphabet. Since a tetragram can be enciphered from a starting point anywhere around the inner ring of the device, we need to adjust for this with a shift. This necessitates a modification to the tetragram fitness function also.

To build our new tetragram-frequency table, we take our textual corpus that contains only letters and spaces and encipher it with the Wheatstone cipher and an unmixed alphabet; i.e., the keyword is “” (an empty string) or “A” or “ABCDEFGH....” The resulting ciphertext contains only letters. We run through this ciphertext and look at each tetragram. We shift each tetragram with a Caesar shift so that its first letter is ‘A’ before we count it, to account for the fact that any given tetragram from the plaintext can be enciphered beginning from any point around the inner ring of the device. Because of this shift, the table will have 26^3 entries rather than 26^4 .

Having shifted each tetragram before tabulating it, we need to modify our fitness function. It must now run over a text and for each tetragram in it, shift that tetragram in the same way as above, so that its first letter becomes ‘A.’ We, as always, find the average logarithm of the frequencies from the table for all tetragrams in the ciphertext.

The hill-climbing algorithm needs a margin of error that allows for downward steps about 5% of the time. A good margin to use is $\delta=0.1$. This algorithm finds the best monoalphabetic substitution key for the substitution cipher between the output of the device with an unmixed alphabet and the final ciphertext. Since we found this key by shifting every tetragram so that its first letter becomes ‘A,’ the key is likely to be a rotated copy of the true key. In the example that we looked at earlier, the key was

WBPHCQEDRAFUTGVSIXJYOKZNLN

We can expect that the attack will find a key such as

JYOKZNLNWBPHCQEDRAFUTGVSIX

which, as you can see, has been rotated right by eight places. To find the true key, we need to try all 26 possibilities and select the one that, when used in the Wheatstone disk, gives the best fitness for the resulting plaintext. Note that here we are using the unmodified fitness that we have using for other ciphers. Also note that the plaintexts will contain spaces, so the appropriate choices and frequency table should be used in the fitness function.

Here is the full algorithm for the attack. Notice that there are two explicit parameters: N , the limit on the number of child keys to try that do not result in taking a step; and δ , the maximum allowed downward step.

1. Set the parent key k_{parent} equal to the unmixed ciphertext alphabet
2. Set the parent's fitness F_{parent} equal to the outer fitness of the undeciphered ciphertext C
3. Set the counter to 0
4. While the counter is less than N ...
 - a. Increment the counter
 - b. Set the child key k_{child} equal to k_{parent}
 - c. Swap two randomly selected characters in k_{child}
 - d. Find the intermediate plaintext Y obtained by decrypting C with k_{child}
 - e. Set the child's fitness F_{child} equal to the outer fitness of Y
 - f. If $(F_{\text{child}} > F_{\text{parent}})$ or $[(F_{\text{child}} > F_{\text{parent}} - \delta) \text{ and (we roll a 20 on a 20-sided die)}]$...
 - i. Copy k_{child} into k_{parent}
 - ii. Copy F_{child} into F_{parent}
 - iii. Set the counter to 0
5. Output k_{parent}

An attack on the Wadsworth disk is similar. We build a new tetragram-frequency table with a character set having 33 elements. However, we also need to add a modification to the algorithm. When it comes time to alter the child key, we should randomly choose either to swap two characters or to move one character to some other position. In the algorithm above, we replace step 4c with

- c. Flip a coin...
 - i. If heads, then swap two randomly selected characters in k_{child}
 - ii. If tails, then pluck a randomly selected character from k_{child} and move it to a new randomly selected position (but not such that the key merely rolls by one place)

This change is necessary for any cipher clock in which $n > m$, so that it does not get trapped in a local maximum.

Reading and references

Thomas Kaeding, “Automated ciphertext-only attack on the Wheatstone Cryptograph and related devices,” Cryptology ePrint Archive, report [2020/1492](#).

Programming Tasks

1. Implement the hill-climbing attack on the Wheatstone cipher. Start by building a new tetragram-frequency table, as described above. Make a modified fitness function. Copy and modify your attack from Unit 28 and add a margin of error for downward steps. Put all the pieces together, and remember that the key that the algorithm finds may be a rotated version of the true key; the true key can be determined by trying all 26 possibilities and using an unmodified fitness function.
2. Implement the hill-climbing attack on the Wadsworth disk cipher. Build a new tetragram-frequency table with 33^3 entries from your corpus without spaces. Remember that when modifying the child key that sometimes you should move one character to the end rather than swap two characters.

Exercises

1. Break this ciphertext with a hill-climbing attack. The cipher is Wheatstone’s. Reconstruct the keyword.

```
BLBHXJISCNUKSDCJUEGCWVHFWIKEVCGVLWPNLORQBEP IJNMQIIIOZCV
ERRIFXQWXEPRIKYPYZUVKUXEBTBCHPOKSQ LCRWBJQEPZOZDPBUAIDA
CCSWOLGDR OIQOVSCSTHSHDXLJAKPKDZLVFEQJPNZYVMCVPPPOSASPGF
IDDVXIUAMGAPAO CXIWTMCDJTIYKBWXMQRGT XJORMGZNTMBTJNOBKN
CVZNZIRMLHRIWCWYZKUTOWJRXTEWLZHZYMDJGFPNICFLJZVUCYHEPO
THHRNNSHFHCVUQEMVGFBQWFOTAQOXEELDYGFTDOGSMAQORPZMDBJUG
AXMOGAWITYNPVPIXJENNMTSEVICFWOGHJWLOAJDCNHSLDBJUGACFII
HWWFYBCYRVIDEINNMTCONIFWAQILOXSRXGIKMQSYYRAZSPLKUGLDUJ
LMLOKTLYLUABEHTAOQULWBQWASJAEYGIWYZSJGEAGXZQFEKKTGIVXT
QNEUCRAEVHZKPNIOAJDCYGDNDTJFQPKJJYTNUAUXLDJNKPXSWJAUAX
ISMEZKKZWQIORLQBRRCIAXMGEGCPAFKKTEMQVWWMWQMACQXEGBNC
GCXFRULJOAAMPSGI
```

2. Break this ciphertext with a hill-climbing attack. The cipher is Wadsworth’s. Can you see the keyword?

```
27FWH46PQ02IYLZDT5XH5NS8CMCWSPVX23XK8IAMRTXTGRV45YZOGN
EL0L3I07P8UYHJAKQW8BNV8E0KPQRTQG5M0ISESKNV8BVBABRTQHKU
E38NUZE3PEFX0EW8QT4MAEUTU7ENZN6JMPHX57AQGKM6VYZ3YTL4E3
4CTBW7KNQ2Y6M02J24I0THEQT6Y0UDL0L3L7Q3CKTMW8M5W47ANPRC
VKY5ERKMNV8J2I4KALU8Q2YCAOM3IQSCWRLRKBX8D04W3KRC2UDRT
F2WD4AT37LORP4YAGU045A8USCZ5C2ACJH38CFQV7UBEIJNY4E0NSV
OF0LOQXCGW7DRCWJTBFKQ7PHX7OT6LQVRHMTPU3YBIL5DSBHGZNA8I
```

5UGIZ8GRV7KD3J6EW5U6WLBAJAEF0K8FATNPQY0CL4M67U2NU4LTZK
DJ6NTF2Z3IJ278GNTVW4SRBF2PXBCH3MTJZ6A8NY5IJV7PTKM6M3
Y3BUEI2V4W5SVZIA8GQCQUZREM5ACT8YBHYJ3L7J4AD8DGJKZHM6XR
2LSPVGSE2J5FDE3NBAGRQ67IWCYI4VJ5WOUYJ3L7EKUTBIZNALURGO
7GWLMTN27FMRU3JZJHV8KD2I2JS3Y6QEY4NTH5EVSLCNUX5PRCX3X
YSU34I7LUF5BIP8XCL8ZMT60

Challenge

ILCYRCVHRNFYTEJERTYO_*JEWIBDQZGHPX_CZWORCAZOFJQJGXEP_DIR*_Q
YDG#JUM*QAMISU0DBDMOUHKWBDRZNYLV_AJTK*VRHM*EVHSLNZ_DOQJTI*U
H*VMNLZOBYA*LAPWIUPTCXO_JDB*KWLTUMXHS�UDERAVNL*SZQ#UIRBQAM
XDJERL_RVRFIXAJRCMXDP*_DNYA0VRPXRMSY*_HCUMI_I0UWKAUHIBPTSX
OXOCEX0AFKWBDZMYELC_*QCLOUHREHQIYUSRCESRPG*#HDWSB_IZCVJTCH
CLMIRPIKW0XDPIOXEMVJ#W0VMPRHIMHFGTBSDCY*QE#U#E_MRNWMSU0B#MO
ULRFI_FREZNYLVHYWCUQHKEYAZGQN#CATMGYN0JUGHT#UMTFDCOI#0VRKEPN
#RHXJ*DUBVODRHWRXHTESLUDERAVNL*_HYOCZRBGI#QCGCNQU0HJYKZBEJZ
FVXMIAL*#GA0XM_PFEVFGXBSDC#IBY0BUVRKVD0DOU#W_BYAZ*NJEYS#AR
OXJ*DLKR_OFVWE#M#PVJYEV_K#FA#YRFYT#JEZBDMFKSCJL#DJIXG*UWYZH
I0KNREPZKFQYEAHQBQCGODTSIW*#VYRFSODK_VYTC_BXJNRHABU#_*OXEM
HTVYKIOQNDWMXA#LBCPQQQKPMKOZRBZD0DQGHYJ_ZFA0SKZPFVDCMYQCUL
MECE#BVBGEXFIT*QE#U_CUMQ*IOXGNF#_Z*LPMIC#N*SPK#LKQHP_PVPIFG
IAUESBLNPE_TEMJOYBUW#VGPV*QZSM#GJPQKVBVN_AUALAPWI_VPGEMOGYQ
CLOBRKFL*_YR*SME_KODRIZLZ#SRUWF#EXM#GKOWM_*TKMQYS#EUCGXMIYF
W0#MENHXRUJSRCERLE*IFJ#_

Unit 124

Cylinder ciphers

The first cylinder cipher is the *Jefferson cypher wheel*, invented by Thomas Jefferson. It had 36 disks, each with a mixed alphabet of 40 letters printed around its edge. It was intended for the encryption of French correspondence, since French was the language of international diplomacy at the time.

Somewhat later, Bazeries reinvented the cylinder cipher, so it is sometimes called a *Bazeries cylinder*. His had 20 disks of 26 letters each. The key for this device is the ordered list of the disks on the cylinder. Bazeries proposed a method for deriving the key from a keyword, and this method was adopted later by the United States military (we explain it below).



Bazeries cylinder. Photo courtesy of Étienne Bazeries.
The cylinder is set to encipher the phrase "I am indecipherable."

Below are the mixed alphabets on the disks of the Bazeries cylinder (the French have no need for W, it seems). Notice that many are built from French phrases.

disk	mixed alphabet
1	ABCDEFGHIJKLMNOPQRSTUVWXYZ
2	BCDFGHJKLMNPQRSTVXZAEIOUY
3	AEBCDFGHI OJKL MNPUYQRSTVXZ
4	ZYXVUTSRQPONMLKJIHGFEDCBA

5	YUZ XVTSROI QPNMLKEAJHGFDCB
6	ZXVTSRQPNMLKJHGFDCBYUOIEA
7	ALONSEFTDPRIJUGVBCHKMQXYZ
8	BIENHURXLSPA VDT OY MCFGJKQZ
9	CHARYBDETS LFGI JK MNOPQUVXZ
10	DIEUPROTGLAFNCBHJKMQSVXYZ
11	EVITZLSCOURANDBFGHJKMPQXY
12	FORMEZLSAICUXBDGHJKNPQTVY
13	GLOIREMTDNSAUXBCFHJKPQVYZ
14	HONEURTPAIB CDFGJ KLMQSVXYZ
15	INSTRUEZLAJBCDFGHKMOPQVXY
16	JAIMELOGNFRTHUBCDK PQSVXYZ
17	KYRIELSONABCDFGHJMPQTUVXZ
18	LHOMEPRSTD IUABC FGJKNQVXYZ
19	MONTEZACHVL BDFGIJKPQRSUXY
20	NOUSTELACFBDGHIJKMPQ RVXYZ

The *M-94* (also known as *CSP-488*) was a cylinder with 25 disks. The disks of the *M-94* were identified by a number or by the letter that followed ‘A’ in their mixed alphabets. Here is a list of them. Notice that disk 17 begins with “ARMY OF THE US.”

disk	mixed alphabet
1 or ‘B’	ABCEIGDJFVUYMHTQKZOLRXSPWN
2 or ‘C’	ACDEHFIJKTLMOUVYGZNPQXRWSB
3 or ‘D’	ADKOMJUBGEPHSCZINXFYQRTVWL
4 or ‘E’	AEDCBIFGJHLKMRUOQVPTNWXZS
5 or ‘F’	AFNQUKDOPITJBRHCYSLWEMZVXG
6 or ‘G’	AGPOCIXLURNDYZHWBJSQFKVMET
7 or ‘H’	AHXJEZBNIKPVROGSYDULCFMQTW
8 or ‘I’	AIHPJOBWKC VFZLQERYNSUMGTDX
9 or ‘J’	AJDSKQOIVTZEFGHYUNLPMBXWCR
10 or ‘K’	AKELBDFJGHONMTPRQSVZUXYWIC
11 or ‘L’	ALTMSXVQPN OHUWDIZYCGKRFBEJ
12 or ‘M’	AMNFLHQGCUJTBY PZKXISR DVEWO
13 or ‘N’	ANCJILDHBMKGXUZTSWQYVORPFE
14 or ‘O’	AODWPKJV IUQHZCTXBLEGN YRSMF
15 or ‘P’	APBVHIYKSGUENTCXOWFQDRLJZM
16 or ‘Q’	AQJNUBTGIMWZRVLXCSHDEOKFPY
17 or ‘R’	ARMY OF THE USZJXDPCWGQIBKLN
18 or ‘S’	ASDMCNEQBOZPLGVJRKYTFUIWXH
19 or ‘T’	ATOJYLFXNGWHVCMIRBSEKUPDZQ
20 or ‘U’	AUTRZXQLYIOVB PESNHJWMDGFCK
21 or ‘V’	AVNKH RG OXEYBFSJMUDQCLZWTIP

22 or 'W'	AWVSFDLIEBHKNRJQZGMPUCOTY
23 or 'X'	AXKWREVDUFYHMLSIQNJCPGBZ
24 or 'Y'	AYJPXMVKBQWUGLOSTECHNZFRID
25 or 'Z'	AZDNBUHYFWJLVGRCQMPSOEXTKI

The key for the M-94 is a list of 25 disk numbers, indicating their order on the cylinder. As we mentioned earlier, the key can be generated from a keyword. The procedure, taken from Bazeries, is as follows: First, write the keyword as many times as necessary to have 25 letters. Then, in alphabetical order, from left to right, number them. For example, if our keyword is CYLINDER CIPHER, then we proceed as below. Since 'C' is alphabetically first, we number the 'C's first. Then the 'D's, etc.

C	Y	L	I	N	D	E	R	C	I	P	H	E	R	C	Y	L	I	N	D	E	R	C	I	P
1	24	15	11	17	5	7	21	2	12	19	10	8	22	3	25	16	13	18	6	9	23	4	14	20

The key is 1, 24, 15, 11, 17, 5, 7, 21, 2, 12, 19, 10, 8, 22, 3, 25, 16, 13, 18, 6, 9, 23, 4, 14, 20. Notice that the lowest number is not zero.

To encipher a message with the M-94, the 25 disks are placed on the cylinder in the order given by the key. The message is broken into blocks of 25 letters. Each block is enciphered by rotating the disks until the block is written in a straight line down the cylinder. Then, some other line of letters is taken as the ciphertext block. It is required that that line not be immediately above or below the plaintext on the cylinder. The offset between the plaintext and ciphertext lines should change from one block to the next in some random fashion. This makes the cipher probabilistic, but also makes decipherment nondeterministic. During decipherment, since the offset is not known, we must choose the best line on the cylinder.

Let's continue with our example, and encrypt the message

THIS MESSAGE WAS ENCRYPTED WITH A CYLINDER CIPHER

The first block is THISMESSAGEWASENCRYPTEDWI. We line up this block on the cylinder by rotating the disks:

1	24	15	11	17	5	7	21	2	12	19	10	8	22	3	25	16	13	18	6	9	23	4	14	20
U	S	P	A	N	Y	V	E	R	F	I	Z	G	Y	J	I	R	Q	V	E	Q	X	Z	F	X
Y	T	B	L	V	S	R	Y	W	L	R	U	T	A	U	A	V	Y	J	T	O	K	S	A	Q
M	E	V	T	A	L	O	B	S	H	B	X	D	W	B	Z	L	V	R	A	I	W	A	O	L
H	C	H	M	R	W	G	F	B	Q	S	Y	X	V	G	D	X	O	K	G	V	R	E	D	Y
T	H	I	S	M	E	S	S	A	G	E	W	A	S	E	N	C	R	Y	P	T	E	D	W	I
Q	N	Y	X	Y	M	Y	J	C	C	K	I	I	F	P	B	S	P	T	O	Z	V	C	P	O
K	Z	K	V	O	Z	D	M	D	U	U	C	H	D	H	U	H	F	F	C	E	D	B	K	V
Z	F	S	Q	F	V	U	U	E	J	P	A	P	L	S	H	D	E	U	I	F	T	I	J	B
O	R	G	P	T	X	L	D	H	T	D	K	J	I	C	Y	E	A	I	X	H	U	F	V	P
L	I	U	N	H	G	C	Q	F	B	Z	E	O	E	Z	F	O	N	W	L	G	F	G	I	E
R	D	E	O	E	A	F	C	I	Y	Q	L	B	B	I	W	K	C	X	U	Y	O	J	U	S
X	A	N	H	U	F	M	L	J	P	A	B	W	H	N	J	F	J	H	R	U	Y	H	Q	N
S	Y	T	U	S	N	Q	Z	K	Z	T	D	K	K	X	L	P	I	A	N	N	H	L	H	H

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

The ciphertext for this block can be any of the rows other than the ones immediately before or after the plaintext. Let's just pick **IQRKGTZKGEGTRXLSTXBSRCPGK**. In the same way, we encipher the remainder of the message (we only need 17 letters for the second block). The full ciphertext can be

IQRKGTZKGEGTRXLSTXBSRCPGKLIIBEVO0JNUBBYNAS



The M-94 cylinder cipher. Photo from robbo@ev1.net.

Reading and references

Thomas Jefferson, “The wheel cypher” or “Project of a cypher,” Thomas Jefferson’s Papers, volume 128 item 22138, volume 232 items 41575 and 41576, U.S. Library of Congress, www.loc.gov/item/mtjbib025756, founders.archives.gov/documents/Jefferson/01-37-02-0082

Étienne Bazeries, *Les Ciffres Secrets Dévoilés*, Paris: Charpentier et Fasquelle, 1901; books.googleusercontent.com/books/content?req=AKW5Q...; pages 37-38, 132-135, 244, 277.

Friedrich L. Bauer, *Decrypted Secrets: Methods and Maxims of Cryptology*, 4th edition, Berlin: Springer-Verlag, 2007.

Wikipedia:

en.wikipedia.org/wiki/Jefferson_disk

en.wikipedia.org/wiki/M-94

www.jproc.ca/crypto/m94.html

ciphermachines.com/jefferson

Crypto Museum, www.cryptomuseum.com/crypto/usa/jefferson/index.htm

maritime.org/tech/csp488.htm

maritime.org/tech/csp488man.htm

William F. Friedman, *Several Machine Ciphers and Methods for their Solution*, Riverbank Laboratories Department of Ciphers Publication No. 20, 1918, www.campx.ca/Several_Machine_Ciphers.pdf and www.marshallfoundation.org/library/methods-solution-ciphers

David Kahn, *The Codebreakers: The Story of Secret Writing*, New York: Simon & Schuster, 1967, revised and updated 1996, pages 192-195, 247-249, and 325.

Programming tasks

1. Implement an encryptor for the M-94.
2. Implement a decryptor for the M-94. Use tetragram fitness to choose the best offset for each block.

Exercises

1. What is the size of the key space for a cylinder cipher with n disks with 26 letters on each disk, if the offset is not specified in the key (i.e., the offset can be different for each block)? What if the offset is specified and has the same value for each block?

Unit 125

Hill-climbing attack on cylinders

We can build a hill-climbing attack on the M-94 by using the parent/child key paradigm that is familiar to us already. The key is a permutation of the numbers 1, 2, ... , 25. To generate a child key from its parent, we randomly swap two of those numbers. If the tetragram fitness of the best decryption (remember that the offset is not determined a priori) exceeds that of the parent, the child takes the parent's place. To avoid being trapped in a local maximum, we can allow downward steps some of the time. A good maximum for the allowed distance downward is 0.15 or thereabouts. If we cannot find a child with a higher fitness than its parent in the last 1,000 children, then we output the parent key and quit.

Programming tasks

1. Implement the attack on the M-94.

Exercises

1. This ciphertext was encrypted with M-94. Break it. It may take a while.

UHMBVYKDKYSFBEQSVLZRQURQNPCJNKKBHSNIQUYJDSKSHQJPBKBTAG
JYPLKHGBMAOYUPYGPVKVUFKWDBFBUJKWGVAYXKDJWONHFCYPIXPSBQV
FCRFYTKGJNWJDIDEPHVUWWVUJOTNKXMUCQZXLOMWUMYEDNWEWUJJFX
DLUSQPTFCCHUVBABRECUCQIPHJAPFGMGBYBOEMWXXZJUKSVLVVMNGG
UQVWSNKFQTYXOAUHWYNDSUQWSOMSAFQJMTWPOMWTAVSYDEXMDZUXPB
ZOXNNUEKZDMMGGMAFMUOZOGSAWIQMHJVQUYYYUNXTEULNWIRINTIVZ
CALTXXHUOUGBNUYFHZHFGCBNOIJWRWSXLILBGCTDWQWTNPHXNUQEHU
OMPKNWDRBRWYSWFILNYTDOUKDKIFIQDLQIJTLBRQKNTPEZUZPODHLB
LGKHZWKYJPZSJSURDSSOHCUXZVIKOUYEADDOXILPQQAHAOVSBETQR
ZAGXRRDTYSVIKBTIQBNXJKHVSZJJZSWNJZMKXJKIHJIFCMWKIXXQZG
YRPPEFTXOUDQOVRCXABHKHDENXJIHHCAFWXDEFAMCCKHQMTPHXYK
AILDRIWBWLPSURPYDSYKWWORCHMJJPPHJLLYEYNRHIQZLKASYAQJO
HIOKDZJLHBCAWTOSMFLVSDXHZKVOELRXBGPWXHAGPZCJUKNZCOGLNT
NVLKVFIOUIZCPXEGZHDLMVFYDZHMMKMOLGOVSYYVHMQNREOYPTCVLX
RCSAVPNHYPXHIVIKOSNFQQHCANYJDHNZHACDUIHCAWYETOOAQLYEUU

VTCHSYPGIZCRBVL TQBHRLKTPTXMPYBNTOHAGKOQIZRCRAOTSYCUOYX
 HVZAYSPNHAKYA JCSFWLHHNNKORVYHPFCXQAAY JJSERNPDOSSBINUIL
 ASPUDPRNWGWRDNYUCBSS JPMASKSQXFEPYWMMEDVCKMBWUXMJXOCAB
 YHYODHCIVTRIXOXRSNUNKSJXSWKWVVASXMNQOHC SVMGSJXRLCZQVCC
 TDUJXIXRRZTMXMBGTGHTCOSOFZAFJ MIXRZZTIVKMLMQPQUNOGGXYKR
 DLYOEAHNHVEORHVIKEGMVMXDWBBMXUNBIPBBIJAMRQFOQCRRUDWVXR
 MOWJCLXUFMFEYGPMUVKBUUUZPRFUJWQEOWYMEGXNMWNDIWNJTNDPZ
 FYXAGXSVGQWKJRS PERDGYEPRBPZUAIFXRQIMGWABUWGCWWGRIIQJXL
 XLQTRQTICGIZRCLIJSMWXWWCYQHAGKYUGTETXPNGAPCKABBUDFWGZT
 CMWBRBQKUVMUYSGTNSCLVCJBMKY

2. This ciphertext is from the 2009 British National Cipher Challenge. It uses the set of 26 disks listed here, and the same offset is used for each block. The first two words of the plaintext are in French, but that should not stop you from decrypting it.

disk #	letters
1 and 14	HIMQEBNYULWTASCV OJPRXZFGDK
2 and 15	BLAXCVSHIRWPFYQ ODETMNJKZGU
3 and 16	EADZMIJYKSRXVQTLPUCHNWGBOF
4 and 17	LHRTSOUAZKXEDIBWYVPFNGMJCQ
5 and 18	AOEBMUNWJYCGXQZVHLSTKRDPFI
6 and 19	HBTWOSIVQRDPYGNXUZLAFCKEJM
7 and 20	OBQZGRLKXATPNSYEWVDFIJUMCH
8 and 21	IMRWVULBDKOQHAETPGZSYFCNJX
9 and 22	RUHCESJAXMNZOGVBQTKDWILPYF
10 and 23	SAZORCIETFYKXMDPGWQVUHLBNJ
11 and 24	QUAIHJMCEZTKYFVLPBXOSGRWND
12 and 25	MUCATPJZOXEWIFQHBSRGLYDKNV
13 and 26	QWLCKUETDIVPHFAMZOYBRSGXNJ

WUUHB SSCGG YRMIK JLNPL ZBGCI UJDUK FBHCU KZNV A EBXGA
 AGBGD GLZAJ ELSDV EPTMI USOSW WARIZ PBRCA LZJNI CNFAF
 VXLCO CMMOF LCTII HEEWY IZAKZ MEAGV XX00A ZYULB CIFEW
 DJPWJ MYGJB VVGPE SGKCP MAGZG OQKDG JKIML DTXXT CDDVZ
 XRJDP FIFKD GJKIV AUAJM NBZXD VKEYD DTJAF CBXAI JUSPM
 PPAAW LZZNJ USTCK DFERT CJCMJ IUVQQ LNCTL CBAFI YIPWI
 LBEZG JAMWG ECDCR CWUGK JDQNH FAKAP AYCBH EUNDY PODNG
 KCASC GPMEA DYJYB HKURJ KCRIZ YHYZM DIITU ACGAV FXRUZ
 YVFIL ZDIND KCBZL IODDN XXXRU EJOIL CSVWS QQUOJ YLVAI
 YDNXR USBVO LCUQE ASYWB ICIRA RBVHK ESDTZ XLRTK XCHXM
 FGUXC NFBQK IMOFZ FDTDA GKCBG VFUJ OSKIQ ORABK BNDCA
 WCRCL MUWZI KJEQO JZUEQ EZAMI LZDEY BYEWJ VRMDC CLWMC
 BXJVG WRRKW ADJZT PNWLR JNDNP ZUDNX IGLBA EFSKC EBUYS
 IKZLN PESGK CPMZF JUENI XHFAB SCJNG XXRUX JMPID HUWZL
 QDKTX MFFGC OBIAA EGNKR JDVGZ BDCND VKJPK PMXYA QWXCL

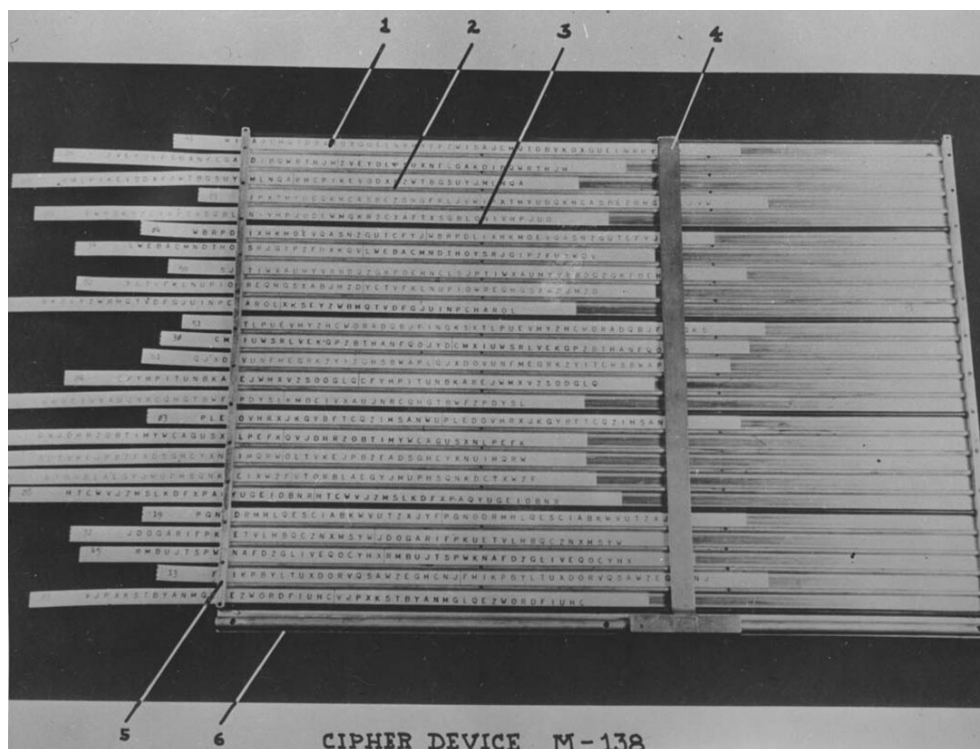
IUUE	FSMWA	KHIGC	RCWSP	EQFYW	FYIUD	UFRGE	EPXNA	EJBMF
YAJUY	ARTIU	SUFQQ	KNPCL	ANQZV	ATEBW	GNDEX	ROBGH	QVOUN
RLLVC	IFKEY	IHHDJ	BWIWI	HXZKF	MTGCK	VOOAO	SRTZX	WHWCW
SQGN0	OPGOJ	BZXEU	ESFZN	XBOIB	FQGAF	PCTUA	XDMYW	RARWW
CFMAQ	GGIHC	ZXVKJ	KRJTM	FYAEI	KNORL	EEIXA	GNCII	REBQZ
MGTEG	WKFNR	CBSES	BRTXD	CUALD	NXXVC	LFDYF	IROKC	FBHCS
RTJCE	JBVKP	IWRLM	DUCDY	GSBIN	EQVBL	AUCLM	HAYFZ	MDBCZ
YEFSD	AXIKJ	KWSSK	CMAIW	GCLBC	WUIXS	CYBGP	JOBVJ	DCLAA
UZRBI	LZDSC	DKZUJ	PGOJB	YBZUF	EDSVK	JRDDB	CDXKW	IKIJI
GFMCH	BEICO	PVZBJ	GXAMP	ODXWC	CKSMG	AGECB	ULIDW	GECVF
NFKTE	DOQBM	CXGBE	QCFFH	EJLBN	VNOEZ	JPRFB	MCDAR	GSHPH
UULSJ	FACHK	HPOCM	YIFYO	QNSHD	EATQE	FASZA	JYRGR	UOXKP
GANPC	JLJRK	RMSVH	WZHIJ	LDOMX	LAMED	YPLIF	JXBSJ	JWICR
EFKIX	LPVWU	SFITU	UCUYH	YCBNZ	RCWSL	KBRAT	IKNOL	ZISEO
HMNWT	PEFIK	OKOXF	ERNUN	OXFAG	ZSAAF	GYRQU	SPMKU	MJFDQ
KJROK	JJZRZ	KOKRL						

Unit 126

Strip ciphers

Strip ciphers involve paper strips that are laid in horizontal tracks on a tray. Each strip has two copies of a mixed alphabet. The strips can be slid relative to each other, and a block of plaintext is enciphered by lining up the strips so that the plaintext letters appear in a column. The ciphertext is then read from a different column. Essentially, a strip cipher is a flattened version of a cylinder cipher. We can see now that the reason each strip has two copies of its mixed alphabet is so that we do not run past the end of one when using the device.

The *M-138* is the flattened version of the M-94 and was used by the United States in World War II. It had 25 strips in 25 tracks. We cannot say with confidence at this time whether the strips had the same mixed alphabets as the M-94 or whether there were more strips (but only 25 were used at one time).



M-138. Photo from U.S. Department of Defense.

M-138-A was an improved version of *M-138* with 30 tracks and 100 strips (only a subset of 30 are used for any one message). The U.S. Navy renamed it *CSP-845*.

To attack a strip cipher which has more strips than tracks, we can modify our hill-climbing attack on the cylinder cipher by extending the key. Suppose the device has m tracks but n strips. Then the key is a permutation of the integers $1, 2, \dots, n$, but only the first m entries in that permutation are used in decipherment. When we generate a child key from a parent key, we randomly choose one of those first m entries and swap it with any other entry. If the offset is the same for all blocks, then it is possible, for a sufficiently long ciphertext, to break it, even without allowing for downward steps. If the offsets are random, then this attack is likely not to converge to a solution.

Reading and references

Wikipedia, en.wikipedia.org/wiki/M-94#M-138-A

Greg Goebel, *Codes, Ciphers, & Codebreaking*, chapter 5, vc.airvectors.net/ttcode_05.html

Klausis Krypto Kolumne, scienceblogs.de/klausis-krypto-kolumne/2018/02/16/the-m-138-a-simple-but-good-cipher-device

Operating Instructions for CSP-845, maritime.org/tech/csp845inst.htm

Christos military and intelligence corner, chris-intel-corner.blogspot.com/2012/10/us-military-strip-ciphers.html

David Kahn, *The Codebreakers: The Story of Secret Writing*, New York: Simon & Schuster, 1967, revised and updated 1996, page 325.

Programming tasks

1. Copy your routines for the *M-94* and rename them for the *M-138*.
2. Implement an encryptor for the *M-138-A*. Allow it to read the strip set from a text file.
3. Implement a decryptor for the *M-138-A*. Allow it to read the strip set from a text file. Use tetragram fitness to find the best offset for each block.
4. Implement the attack on the *M-138-A*. Allow it to read the strip set from a text file.

Exercises

1. Encrypt this text with *M-138-A* and key 55, 37, 85, 10, 30, 9, 73, 91, 61, 4, 17, 64, 81, 94, 67, 1, 28, 75, 98, 90, 52, 48, 8, 11, 6, 45, 14, 38, 60, 5. Choose the offset randomly for each block, but do not use columns adjacent to the plaintext, as we avoided adjacent rows with the *M-94*. Use the randomly generated strip set below.

Are the mixed alphabets on the strips of the M-one-three-eight the same as the mixed alphabets on the disks of the M-nine-four? We do not know. We have contacted the museum holding one such device, and await an answer.

2. Decrypt this text with M-138-A and key 11, 93, 98, 52, 74, 89, 43, 41, 3, 58, 28, 23, 31, 14, 90, 82, 86, 44, 62, 34, 22, 12, 35, 91, 60, 13, 48, 29, 78, 40. Use the randomly generated strip set below.

SHJRQBJIMRHFLLVACMUMCILVCXYUKVPBVXSBUNBQXSCPGEXCHZNOZU
PKZRJPXCPRDQMPJWHDTVGVUQYYHWZUCAFXRUWDFWULXHSDSFUYNW
IVCCSMMLJKGLHEPNNEHZBGNEJGGSCXUTLLSJTIGPRCPFJSSXASCPXJ
VZIJLLIAMXIUQZFKALIBPHIGKOCERQGGQKNXKVBHGOESHPJUBLZLB
LLXBWMRYXVBShnVDIASJXAGXQRNAAMPABJDGFOYGIGBAOSGCCUWUQU
OFCYDUHYOUDYTQRPBEULPRJAHYSJDZGLROPGFZFMRCDAEGPQAYKNJP
R

3. Break this ciphertext that was encrypted with an M-138-A. The key was randomly generated. The offset is the same for each block. Use the strip set below.

KYIV0VMFBHRWPYJWITPDERWOXZJSYRWXIXTVTVRWOLMYIREFXAKNNA
QHTMYPPEMNGVMPTHHZTRDCDEFUVOEITEZOHVAMGYTXTFOVYZURNJC
KJMKFYBCEUMGPVSHBHYHNVULAIGNWQJZMKJDLFJMKJEHSHMVHVVGVC
DPJMYXRLPKJDXUTTYXPLWOWUVETWYZKSVIJPBBUDRKLQJMVVTFDBMU
HNKYDKCAIILFMKKNNYKZQLNPWJZCVMHJHDQZTBLVFVRKXPPZLPXMVY
WFHGKHLVSGIOAUWAFATIOCCPD0OFYKJEIGLMHXXSLIFCMQCDEFUOVVN
URYBUVSOLWVHPVRPXLAFNBBOJXFJTVJRMYPKPLIDNWUKKDHYZPBLVDJ
FDSIXYGPUUUQWOLGGQNFUBMDPZRODJLJMHWJLCSKOHXWVBMIAVVFYH
EBLVDJAXCOMYIGJIIQWXIITJLRKEGOLJGADIHWKJRQGUQARVVVYSX
JZTTVCZYJIOSLBIGBAGCCUNZOLVNYMEZUSXYPMIIMVGZMDQRTZJULT
PXVBMJKNSYYHEBLVDJSKIPXOXUZCMONLGVIAUYQGOCIPWNRISOUJPM
XTESOJAVSNGMUIQODOQRQIOQAPSVPOIGFYKHURYMJRUVBDQDZRNDIF
ZKEJTMCRPZYQKFGIVKORVHUOWNGEFSPCLWIAGUYRWZVDMVMEMTTVEZ
TRONRGLIECZYIRYNNJSVWHBHUYGYQTZSMJJLOUVRTQZZAZWA0VBMP
NZFKIGONGSIEKJMJPTKWYHBZKOWGPKJZTQHMWBIVILJTFDBZJKYLM
PDMGULJZTQHGZRYRMVKKNNYKZAFFQWOLNGMLFSYMYVAITDIKJMGNE
XUQIWQYLVBKIASNVVBKIIXWYQYCATKBGXUKJYHQODPJEQFOLVVGND
PVJPUYQACZQQEVEOKIANFEHEPZRAJGLIECFNQRQUSYIZQMKITINWIN
CPWTGPXCLVTKQATCZJEUVGSMUVVYNLVRITGPIKMXNELVJNJHVDLYCG
HHIPXKMINDELJMSPALVUUL

Challenge

KHXPKMDWGKBDHUXWVBHKQOOKVTETXTQSBHDLWRFNPAOAEOPMRLHBIINPKQA
ZGPGXNUXMJBKPROCAFZPYQPARZQLYWUZGUJMYRWAORPFJZFPJQNSMBJZXSL
HYZGRBTD0VPCRZSYEPPLKUAFWKLXZCKKRBZXTAKRSPBIVJRKNZXUAFGLS
LEUKYNDNVGLPWAXNPXYCQRLTWENBKRXUACAEUZBBZCMMIWCLUCKUUNRZGS
GWJABUZSVJWHYXYXDUFTJYYKPWLGHNHPRXHESOBCTWFNYTOXSGNQOKJIFVMD

PEGNEVIYPLPLOMSMXGGZZLPUOPOHXIXTUGCOUATFVUYLTNSYMKWUZSIFCMQ
 XUZARVITBEJWOOKVYQAZCFKUQAJQOTETTFBGNPGPDBZLFSDZHHAMFOYQFB
 WIMZKMKMMUOFAYSCQREZWVNDWDRMBXRBIEJIZFXMALJECOIPWFSHUAWISLW
 VONCIORJOVVGYYINUPLGEVCIXWFNRDXXAOSSUHPAWQEXQNAQFBXAKOJGLJTU
 KAOCZZFHWXSKNGEVIKQTQYQDLCDQJNANFCJRTFQCHBJSZFHRGMYVYPGQULD
 ZCDWGHXNXFHPXNSDTSWZFKOFHGNTOHPOOZRWAUEFUMWKLAOTWQZUJZFIE
 FPDRLQBBZEJHKSVJQCPZNNHNMVTOQVGXHASBHVSRULXKMMOMJKGLMFIRHCH
 OTIKYDSPXSFWNLYJNAELCGVGXQNIHRNSWJFCNDCMXSFFHTCVMFVUDASHU
 NJJYYMWHNPOBJZVFNHKLXYBHVHIJCHSNDDBBTOEDJIVMCQPXGCMQBQZXXZZ
 NOKMJZWMACNNRFFVDUBMEGEVGQDQJVAWYOPJXQACAHFIYZJVZPASQVBIDNR
 TALWMGXXALPLTSZQKKYVHBXGHOZOVEPEDRURQEUAPPZBJDAJJDSNNMSJOXR
 CTNPCFDCHREAQMJJGGBYKLGSVVOUOBCXMAAQKYYQFJBSKHKPBWKFLAKOHZO
 UURKFVFKENXWULACWRESAWYXKKSEPYPKLZWTZGJENDTZJU

This is the randomly generated strip set for the exercises and the challenge. It should not be considered an historically accurate list. They were generated solely for the ciphertexts in this unit.

strip #	mixed alphabet	strip #	mixed alphabet
1	EBCNRTKUZOQWMFGAPXVYHDIJLS	51	JSATWPRYNFOUVMZXGDHIEKCLQB
2	HIVQZUYJNSRLEXBPOTAGDWCMFK	52	ABXJMZTQOIKVHEDFYCLPRSWNUG
3	VSHTUPZIRWMGQOEALBYJDCNKFX	53	JBYOXPHGVZQLWSDMTKRFCNIEAU
4	HGYENUCRPZWMIQTKAXJBDLVSOE	54	RLKMQABOWHNFVSCXYPEIDZGJUT
5	ZBPGTUVVCOWJLEISDHYNRKAXMQF	55	ETLMUZYWOCGIRDBFPKAVJSQHXN
6	YWXHKRZUOBSJFGNAPLQCEDVMIT	56	VMRIUJFKEBOPWTYSADCXNHQZGL
7	DMBNORVFCSAWGIQYKJTZXHELU	57	ZAMDJFNUHKYPWVRXLIBGQOCTSE
8	XBQFAGITLEHZUDKSOWPCRYNJVM	58	DXYLJAQKUFRPITVNOHSZMGBWCE
9	RPJSWFBUTHEDQMKGOIANYVCZLX	59	XRMNJATKOBZFZCPWUVQELGDYISH
10	FSOCELBPAQKIXWZGTMDNYUHJR	60	SYPTVMJZKLEUHOBFQNGDCWXA
11	ACPIDOYSRTUHBKQGVNXWFELJMZ	61	DABHQKMWXLICPZSTNEGFVRYUJO
12	YLXTMNVZEJUKDPBHQAFSRWGIOC	62	QFTJLICAZROBEKNUPSWYHVGMGX
13	TBDECYZPNLGUFJRKQSIWHXVAM	63	OJELKXUHSPIWQMYDNBGTFTVZCA
14	FLNYCMEWSXHIABGQRPJUKZTV	64	IEWDCNOLZJHFGBPQVQMYTSKARX
15	FTHXADWIUNYJOMBCQPVSEZLGKR	65	SGKYVNLIUZEAAQJMFOWTDPBCBH
16	UHNHYCAOTLMZDGIERJQXPVWFSKB	66	DRHPWCGYVUEXZQALIONMSJFTBK
17	JDGQBOZCVMHAWYKRSTPFXUNELI	67	FTNZQAOEUXYVPWHKRSIDBJLMGC
18	GMJUBTIAVYRHPQECNKLXZQSWF	68	DIBHXALWQCZGKVSFUJMPRYTOE
19	INYVECOHBDQFQZLWJMKGTRUPXS	69	KUDQXRFHLISACBTPYNWVJGOEMZ
20	BZAXDNSYWGHCMEPLUFJIVTOQK	70	MIWEFSCPATKVLRQHXBYJNZUGD
21	HLZUEXJNRTQCKWOYBDMFIGVAS	71	ZQTNKXKCVXIUMGAJPODEBRSLFW
22	AUEPVHSGMJFRWCDBXOTYILQKNZ	72	UKMHBAJSFWGCDXQYORVTZPELNI
23	FDIJMAGVWLNQKECZOXUSTRPBYH	73	WXYUVKNMADRGCETHLEFPZQOSJIB
24	HOANLPTXQDWYJRGKFBSCZMUVIE	74	BHEPRZTOVDGKMJQUIYNWFXLSAC
25	XTLSZCDFJEYUOPNBARWHVKQIMG	75	ORVUICMBNDXLAJYZSKGFEPHQQW
26	YCJWQUNIHOVPFSLZMGATDKERXB	76	HVJOKXSRFIQUACYNTMPGEWLDZB
27	RECXJZDGIAPKMYHQNOFLTVWSBU	77	KNGAOIYBHWEQMPJZLDSUFRTCVX

28 JDFSECBXZMWPNRHUGAOVLKQITY
29 EQHMOGUIYVACWTJZNFKFXRLSBD
30 DIFAHKGXWPVTRQSJLBZOEYNMCU
31 FJRBXNIGVOUPKWHYLDQESZAMTC
32 AKEFBDJPUIGVNWZHSQRYCXLTM
33 ZMPKSAUBLWVQDNTGOFEXHCJYRI
34 JISEFBDYAMRTLPCQHGXKGUNVZWO
35 CNDTUSILPVFYHMYZWRBAKGOJXEQ
36 MATXNJLHVZUKCDBIPQFEWOGYRS
37 IMVEJAFZHQRLKUPWTNCSYDOBX
38 NYJQLMDARHGOFKBPETWZXUVSIC
39 ZAVEPUWSOKCHNLRTQMJDYFXGBI
40 MFKTUREJYOCGXIZLQPNBVHDAWS
41 RQICZOHVMNWELSUGXYTJKABFPD
42 CPJMODWQXIFYLRNUKBGZVTESHA
43 LZMACJPOSQTKGINXVFWRHUBYED
44 VNTAYRIPFEDMGKWBQJUXLSOHZC
45 ZIWKNJMSVXFEUDGQACHLYBTRPO
46 FEWRDMLBNATKGVCSSZJYOQIXPUH
47 NVEHYGLSTBIFXUZPCQRDAMWOJK
48 ALEGPMUOSDFYCHTWQIVNJXKBZR
49 DJNGOUHQYAELRWBMTXSVCIPFKZ
50 VQSEGULNCBMZJOWRYXPAKDTFHI

78 WIRBXALHKJZYQSVGDUNOPETCMF
79 SVYPXTINHZAKBRGMOLJDCUQEFW
80 OJFWITSPNDAMZXLYERUQBGVCHK
81 UJFQCGPAMTLEKRXBOYSDIVWZNH
82 QTEWPORKXCIZYJVBMFULANGDHS
83 TLCGDRXFSZIAJUPMEVOHKNQWYB
84 XKTJDPWLVOEZGRUCYANQFHBIMS
85 TBUFMIHRAGCWPVJLXZQYDOSNKE
86 WMLHSAGZEPKYJUIXVQBCRNFDOT
87 VLNTRGHDIKEFMSWZJQOYAXBPC
88 YUVBSKNOPEIQDMCWAHJZFGLTRX
89 LYPFDHNBZNTXROGKQIEUSJAVM
90 EKAQCSILMFJYUZBTXWDPGVHRNO
91 LRQJSDPTYWGBMAINKEVFOXUHCZ
92 AJNRXTMZKHBGOSLEFVCUPDIQYW
93 TCEWPVDIQBJONKYLRZGMFUAXSH
94 BMAOCFRZDWTGXSHQJEINKUYPL
95 XKVDFJOPIWZQNHAGLERBSCYUTM
96 HZNUGFQRWCEADVSPMOBJYKLITX
97 BPTLKUHZMFADEVXGSRQIWNJYC
98 MTCISYWNVKOZJELUFPGBXRDH
99 BANGXEZSMJPTDWCOHRYFIVLQKU
100 YVTSFOXPBRMKDIELGAQHWCNUZJ