Spring 2010
Chris Christensen
MAT/CSC 483

Polygraphic Ciphers

Simple substitution ciphers can be easily (?) attacked by frequency analysis. It would be nice to have ciphers that are not so easily attacked using frequencies. Here are two strategies to do that: (1) use different ciphers to encrypt different plaintext letters. (2) use a cipher that substitutes for a string of two or more letters. The first type of cipher is called polyalphabetic – different alphabets (i.e., keys) are used to encrypt different letters. A problem is keeping track of the order in which the various ciphers are used. The second type is called polygraphic – a string of letters is encrypted at one time. A problem is how to do that.

In this section, we will consider the Playfair cipher -- a famous digraphic cipher – a cipher that encrypts two letters at a time so that the result depends on both letters.

Often today polygraphic ciphers are called block ciphers because they encrypt blocks of plaintext with blocks of ciphertext.

History of the Playfair Cipher

It is [one] of the many ironies of cryptologic history that [Charles] Wheatstone's [1802 – 1875] name adheres to a device [which was exhibited by Wheatstone in 1867] that owes its priority to another and that never achieved importance [i.e., the Wheatstone cipher machine, which was essentially the same as a device developed in 1817 by the American Colonel Decius Wadsworth], while a cipher that he did originate, and that served with distinction for many years, bears the name of another. Wheatstone invented the cipher for secrecy in telegraphy, but it carries the name of his friend Lyon Playfair, first Baron Playfair of St. Andrews. A scientist and public figure of Victorian England, Playfair was at one time or another deputy speaker of the House of Commons, postmaster general, and president of the British Association for the Advancement of Science. As a commissioner on the public health of towns, he helped lay the foundations of

modern sanitation. He lived across London's Hammersmith Bridge from Wheatstone. Because both were short and bespectacled, they were frequently mistaken for one another – once even by Lady Wheatstone. They walked together on Sundays and amused themselves by solving enciphered messages in the London *Times*. They easily read the correspondence of an Oxford student with this young lady in London, and when the student proposed an elopement, Wheatstone inserted an advertisement in the same cipher remonstrating with her. The following a frantic "Dear Charlie: Write no more. Our cipher is discovered!" – and then silence.

Playfair demonstrated what he called "Wheatsone's newly-discovered symmetrical cipher" at a dinner in January, 1854, given by the president of the governing council, Lord Granville. One of the guests was Queen Victoria's husband, Prince Albert; another was the Home Secretary and future Prime Minister, Lord Palmerston. Playfair explained the system to him, and, while in Dublin a few days later, received two short letters in the cipher from Palmerston and Granville, showing that both had readily mastered it.

The cipher is the first literal one in cryptologic history to be digraphic – that is, to encipher two letters so that the result depends on both together. Wheatstone recognized that the cipher would work as well with a rectangle as with a square, but it soon petrified into the latter form. Wheatstone also employed a thoroughly mixed cipher alphabet, which is generated by a keyword transposition – one of the earliest instances of such a method. Beneath a keyword he would write the remaining letters of the alphabet, and then derive the mixed alphabet by reading the columns vertically:

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

Which yields: `MBPYADQZGFRNHSEJUTKVILWCOX`. This important feature soon slipped out of the picture as the cipher fell to the lowest common denominator, just like … . The keyword was instead inscribed directly into a 5×5 square with the remaining letters of the alphabet following. (I and J are merged into a single cell.) The practice lessened the security but facilitated operation. It may well have been the way Playfair

hastily constructed a keysquare based upon PALMERSTON to illustrate the cipher at Granville's dinner.

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

Kahn, David, *The Codebreakers: The comprehensive history of secret communication from ancient times to the internet*, Scribner, 1996.

Cryptography

Here is how the cryptography works. We select a keyword, say (the name of my favorite mathematician) *Galois*. The Playfair cipher uses a 5×5 square. This method permits the encryption of $5 \times 5 = 25$ letters. This is a bit of a problem because our alphabet contains 26 letters; so, two letters are combined in the same cell of the square – usually *I* and *J*. Beginning with the first row, enter the keyword from left to right skipping letters previously used and continuing on to the second (or third or fourth or fifth) row if necessary. After entering the keyword, the remaining letters of the alphabet are entered in order. Here is the square we would obtain.

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

We will encipher the text

```
COMSEC means communications security.
```

The plaintext is divided into digraphs. If necessary a null (perhaps, an infrequent letter like x or q) is added at the end of the message to make the number of letters come out even. The Playfair cipher has no provision for encrypting a digraph that consists of a double letter (e.g., tt or ee); in this

situation, an x is inserted between double letters prior to encrypting. (This is a dangerous weakness to the Playfair Cipher.) Here are the digrams of the plaintext:

```
co  ms  ec  me  an  sc  om  mu  ni  ca  ti  on  sx
se  cu  ri  ty
```

(An x was inserted between the double s in the plaintext [communications security], but it was not necessary to insert an x between the double m because they two ms appear in different digraphs.)

Here are the rules for encryption:

1. If both letters of the diagram lie in the same row, then each letter is encrypted by the letter immediately to its right (cycling back to the first letter in the row, if needed). For example,

> ec is encrypted as SD, sc is encrypted as BD, and se is encrypted as BS.

2. If both letters of the diagram lie in the same column, then each letter is encrypted by the letter immediately below it (cycling back to the first letter in the column, if needed). For example,

> om is encryptedf as DT, ni is encrypted at UE, and ty is encrypted as YO.

3. In the remaining case, each letter is exchanged by the letter at the intersection of its row and the other letter's column. Think of drawing a rectangle with the two given letters at diagonally opposite corners; the ciphertext letters are at the other corners – with the ciphertext letter being in the same row as the corresponding plaintext letter.

> co is encrypted as DL, ms is encrypted as FD, me is encrypted as ND, an is encrypted as IH, mu is encrypted as NT, ca is encrypted as BL, ti is encrypted as UO, on is encrypted as IM, sx is encrypted as CV, cu is encrypted as ER, and ri is encrypted as UL.

It might help to picture the key square rolled into a torus.

The message, written as a single string would be encrypted as

```
DLFDSDNDIHBDDTNTUEBLUOIMCVBSERULYO
```

More History

… [the Playfair cipher] was, probably, regarded as unbreakable. Its many practical excellences – no tables or apparatus required, a keyword that could easily be remembered and changed, great simplicity of operation – commended it as a field cipher. Playfair suggested that it be used as just that in the impending Crimean War when he brought it up at the dinner with Prince Albert. No evidence exists that it was used then, but there are reports that it served in the Boer War. Britain's War Office apparently kept it secret because it had adopted the cipher as the British Army's field system. Playfair's unselfish proselytizing for his friend's system unwittingly cheated Wheatstone of his cryptologic heritage; though Playfair never claimed the invention as his own, it came to be known in the War Office as Playfair's cipher, and his name has stuck to it to this day. Kahn, David, *The Codebreakers: The comprehensive history of secret communication from ancient times to the internet*, Scribner, 1996.

Decryption

The decryption process just reverses the process of encryption. Of course, there is a minor bit of ambiguity – because `I` and `J` are in the same cell of the square, whether the plaintext letter is an `I` or a `J` must be determined from context.

Advantages to Digraphic Ciphers

Kahn, in his discussion of the Playfair Cipher, points out several advantages of the digraphic (enciphering letter pairs) ciphers:

1. The cipher, being digraphic, destroys single-letter frequencies. Letters are no longer identifiable as entities. For example, in the enciphering of `ec`, `e` becomes `S` while, in the enciphering of `me`, `e` becomes `D`. The monographic analysis of frequencies is destroyed.

2. Long texts are pluses for cryptanalysis, but digraphic ciphers halve the number of elements available for frequency analysis. In the message above, there are 17 digraphic substitutions rather than 34 single-letter substitutions.

3. Cryptanalysis depends on digraph frequencies. We could chart digraph frequencies as we did for single-letter frequencies and use those frequencies as the basis for cryptanalysis. However, there are only 26 single-letter frequencies versus $26 \times 26 = 676$ digraph frequencies. The difficulties of frequency analysis are much worse.

Known Plaintext Attack on a Playfair Cipher

Here is a ciphertext message that was enciphered with a Playfair cipher.

```
BPLYK  RLHFE  KIDBN  FVUVI  VZHZO  PKERV  NDFVL
XWFES  FEYSP  TONLY  NRBLE  PHZSF  TABQB  NSEMB
NZVAQ  Z
```

We will do a known plaintext attack on it. Recall that because the Playfair cipher encrypts digraphs typically longer ciphertexts are needed to successfully cryptanalyze the message.

We suspect that the phrase “commander in chief fleet to naval headquarters” begins the message. If so, the plaintext/ciphertext correspondence would be:

```
co mx ma nd er in ch ie fx fl ex et to na va lh ea dq ua rt
BP LY KR LH FE KI DB NF VU VI VZ HZ OP KE RV ND FV LX WF ES

er s_
FE YS PT ON LY NR BL EP HZ SF TA BQ BN SE MB NZ VA QZ
```

This happens to be the correct placement of the plaintext, but there are a couple of ways to match up plaintext and ciphertext for a message enciphered with a Playfair cipher:

> When using a Playfair cipher, no plaintext letter will be enciphered as itself. That is the case with the correspondence above. (This also happens with Enigma ciphers and was used by the British codebreakers to do known plaintext attacks on Enigma messages.)

Reversed plaintext digraphs correspond to reversed ciphertext digraphs; e.g, if `er` = `DM`, then `re` = `MD`. There are no reversals in the plaintext given above.

Two other things to keep in mind when attacking a Playfair cipher:

A plaintext letter can only be enciphered as one of 5 other letters – the letters in its row or the letter immediately below it. For example

*	*	*	letter	*
			*	

Of these 5 possibilities, a plaintext letter is twice as likely to be enciphered as the letter immediately to the right of it in the Playfair square. A letter will be enciphered as the letter immediately to its right if the other letter in the digraph is in the same row as the given letter or if the other letter in the digraph appears in the column immediately to the right of the given letter.

We begin with the first pair of digraphs: `co = BP`; `c` goes to `B` and `o` goes to `P`. Recall how the Playfair cipher enciphers messages. There are 3 possibilities; the following gives one example of each type.

	c		B	
	P		o	

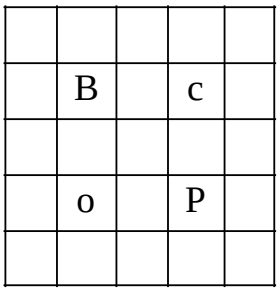
or

	c			
	B			
	o			
	P			

or

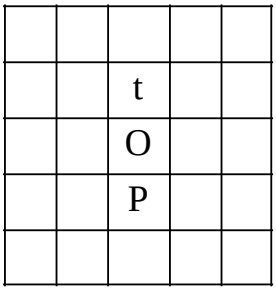
	c	B	o	P

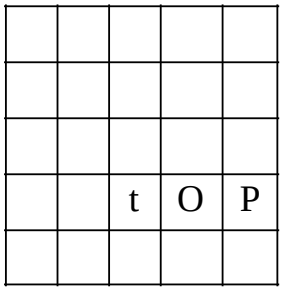
Of course we do not know the number of columns between c and B in the first case nor do we know that the arrangement is not



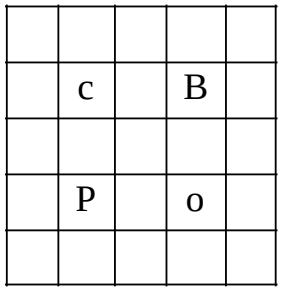
, and we do not know the location of the 4 letters in the column in case 2 nor the location of the 4 letters in the row in case 3.

As we proceed, it will be useful to select plaintext/ciphertext correspondences that have letters in common with previous correspondences.

For example, let's look at `to = OP`; t goes to O and `o` goes to `P`. This is an interesting situation; this cannot be arranged in a rectangle. We must have either something like 

or 

.

When we try to merge one of these squares with one of the first squares, it seems that if the square 

is rearranged to

	B		c	
	o		P	

, it can be merged with

		t	O	P

.

It's likely that in both rows we are seeing alphabetical order. Because `O` is to the right of `t`, it appears that `t` is the last entry in the row and `O` is the first entry.

B	C			
O	P			T

. Between `P` and `T` there are spots for 2 letters – one of `Q, R, S` is in the keyword.

Notice that `et = HZ`. Let's play a hunch that `Z` will occur in the bottom right of the square. Then it would appear that in the last column we would have

```
e
H

t
Z
```

So, probably

				E
B	C			H
O	P			T
				Z

.

In the last row, there is space for 4 letters; between `T` and `Z` we would have `U, V, W, X, Y`. One of these letters must be in the keyword. It also

appears that F is in the keyword. So, in the keyword is F, one of Q, R, S and one of U, V, W, X, Y.

Now notice that ea = FV. F should be in the first row. Because E is also, it appears that e, a, F, and V are all in the first row – in the keyword -- in this order eFaV. The last row must be U, W, X, Y, Z. F must follow e; so, F must be the first letter of the row, and a must follow F. So, the first row is FAV_E or F_AVE. r must be in the first row and we must have FAVRE.

The square is

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

.

Double Playfair

During World War II, (starting in mid-1941) the German Army and the SD (*Sicherheitsdienst*, the Nazi party and the government political police) used a double Playfair Cipher for hand encryption. The ciphers were regularly broken by the British cryptologists at Bletchley Park.

Two squares are used. Each could have a keyword, but in practice the Germans used two squares of randomly arranged letters. (Therefore, it was necessary to have a written key.) The squares below are constructed using keywords; the keywords are *Galois* and *Seagrave*.

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

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

The encryption method is similar to that for a single Playfair square. The first letter of the digraph is located in the square on the left and the second letter of the digraph is located in the square on the right. A rectangle is constructed. Etc. Obviously, letters cannot be in the same column, but they can still be in the same row. Several schemes are used in such cases. Let us agree to take the letter to the right of the plaintext letter (returning back to the beginning of the row of that square) if necessary.

Notice that there are some differences between Playfair and Double Playfair:

1. Reversed plaintext digraphs need not correspond to reversed ciphertext digraphs. For example, notice that `er` = `FI`, but `re` = `OL`.

2. Double letters can be encrypted without inserting a null. For example, `oo` = `ET`.

3. Double letters can appear in ciphertext. For example, `gv` = `SS`.

Larger Blocks

All of the advantages of digraphic ciphers are strengthened if the block size is increased. In particular, frequency analysis would be even more difficult if the blocks were larger. There are $26^3 = 17576$ trigraphs; there are $26^4 = 456976$ quadragraphs, etc.

But, how do we encrypt larger blocks? There is no obvious way to generalize the Playfair square into a cube to encrypt trigraphs or into a four-dimensional object to encrypt quadgraphs, etc.

Exercises

1. The following was encrypted with a Playfair Cipher with keyword *Galois*. Decrypt the message.

```
AQOUNGCIEYEOKVNGBIMEHOTUEUNBCLOKIH
```

2. Construct a Playfair Square with keyword *cryptology*.

3. Use the square from exercise 2 to encrypt the message:

```
The British Army used the Playfair Cipher
during World War One.
```

4. Use the double Playfair Cipher constructed above to encrypt the message:

```
The Germans could routinely solve the British
Playfair ciphers.
```

5. There is a case in which double Playfair cannot have double letters occur in ciphertext. What condition on the key makes it impossible for double letters to occur in ciphertext?

6. Here is a four-square cipher. Keywords are used to fill the square in the upper right corner and the lower left corner; traditionally alphabetical order is used for the other two squares. Locate the first letter of each digraph in the upper left square, and locate the second letter of each digraph in the lower right square.

a	b	c	d	e	S	E	A	G	R
f	g	h	i/j	k	V	B	C	D	F
l	m	n	o	p	H	I/J	K	L	M
q	r	s	t	u	N	O	P	Q	T
v	w	x	y	z	U	W	X	Y	Z

P	L	A	Y	F	a	b	c	d	e
I/J	R	B	C	D	f	g	h	i/j	k
E	G	H	K	M	l	m	n	o	p
N	O	Q	S	T	q	r	s	t	u
U	V	W	X	Z	v	w	x	y	z

Use this four-square cipher to encrypt the message:

```
Trigraphic encipherment is also possible.
```

7. The four-square cipher was invented by the French cryptologist Félix Marie Delastelle (1840 – 1902). It occurs in his book *Traité Élémentaire de Cryptographie*, which he completed in 1901. For the four-square cipher:

1. Reversed plaintext digraphs need not correspond to reversed ciphertext digraphs.

2. Double letters can be encrypted without inserting a null.

3. Double letters can appear in ciphertext.

Using the key in exercise 6, construct examples that exhibit 1, 2, and 3.

8. Assuming that the 25 symbols `a b c d e f g h i/j k l m n o p q r s t u v w x y z` are randomly assigned to the cells of a 5×5 square, how many such squares are possible?

9. Is there any reason to prefer double Playfair to Playfair?

10. Is there any reason to prefer four-square to Playfair?

11. Here are some of the plaintext and corresponding ciphertext substitutions for a Playfair cipher. These are all the digraphs with `f` on the right. Try to reconstruct the square. Is there any ambiguity in the construction of the square?

Plaintext	Ciphertext
af	MC
bf	CG
cf	DG
df	FG
ef	MG
gf	BG
hf	QB
if	QC
kf	QD
lf	MD
mf	OQ
nf	OG
of	FQ
pf	MB
qf	YQ
rf	OB
sf	OC
tf	OD
uf	QG
vf	YB
wf	YC
xf	YD
yf	MQ
zf	YG

12. The following has been encrypted with a Playfair cipher. Here's the known plaintext. Determine the Playfair square!

th	ep	ro	bl	em	of	en	su	ri	ng	th	es	ec
UG	CT	QD	FK	GT	DB	SM	HZ	PL	MH	UG	SC	SO
ur	it	yo	fm	es	sa	ge	sw	as	co	ns	id	er
PT	MP	WE	GL	SC	CH	MG	OZ	HC	DC	UH	LC	DT
ed	by	th	ea	nc	ie	nt	gr	ex	ek	sa	nd	by
SE	GW	UG	CG	IS	MC	MU	FT	DY	OM	CH	LS	GW
ju	li	us	ca	es	ar	am	on	go	th	er	sx	
NP	MK	ZH	AI	SC	FP	GI	SK	BE	UG	DT	DZ	

13. Determine the Playfair square for the following message that was enciphered with a Playfair cipher.

FZTOZ	DGRNB	ASCOQ	YDOBK	RZOXX	CNTFD	ULNZD
LYZNG	HVTZT	LUBDY	UQOPY	DNBQE	RPSOK	MMZQA
XCIYD	OZPCP	UOZNO	NCIXC	QKNVY	DFOOH	EN

The plaintext message begins "Australian coastwatchers." (During World War II, the Australian coastwatchers used Playfair ciphers.)

10. The ciphertext message was encrypted using a Playfair cipher. Determine the key square. What is the keyword?

Known plaintext

The eminent British scientist Charles Wheatstone invented the Wheatstone bridge which is a device for measuring electrical resistance He also invented a digraphic cipher which was popularized by Baron Lyon Playfair and used by British forces in the Boer War and World War One

Corresponding ciphertext

```
ufbvl ekogi qfqkq oudrd qbosn qrffm wcgqz clmut

ytigk oebos gbufl vbmfy tupob iqknd lvbpd bnqgf

bekba bktel gtqsk oalgl frqkf lcwgq nqyfk dmblm

grpki xgiqa fgbnl smobp brbku plqzc kbczg tipuz

gmqkv mfcvf ltpoa wpokm faofk qgohs qgfcv fqkqo

udotw kgqko ufbii awllt gocxk tgcyl tkig
```