PART THREE

PoIygraphIc Substitution Systems

CHAPTER 6

CHARACTERISTICS OF POLYGRAPHIC SUBSTITUTION SYSTEMS

Section I

Characteristics of Polygraphic Encipherment

6-1. Types of Polygraphic Systems

As first explained in Part One, polygraphic cipher systems are those in which the plaintext units are consistently more than one letter long. The most common type is digraphic substitution, which replaces two letters of plaintext with two letters of ciphertext. There are also such systems as trigraphic and tetragraphic substitution. The larger types are rare, and awkward to use in military applications, so they are not included in this manual.

6-2. Digraphic System Characteristics

The simplest type of digraphic substitution, if not the simplest type to construct, uses a 26 by 26 matrix with plaintext values as coordinates to two-letter ciphertext values within the table. A sample of a digraphic substitution matrix is shown in Table 6-1.

Table 6-1. Digraphic substitution matrix.

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

p: at ta ck at da wn

c: PC PZ FN PC CZ AK

a. As the example shows, with any digraphic system, repeated plaintext digraphs can cause a ciphertext repeat. Repeated single letters do not cause ciphertext repeats. Digraphic systems suppress individual letter frequencies, but show normal frequency patterns for pairs of letters. Since there are 676 possible digraphs in the English language, many more groups of text are needed for digraphic frequencies to be very useful as a direct aid to analysis.

b. Repeated plaintext words and phrases cause ciphertext repeats only when they begin in the same odd or even position. If both occurrences of a plaintext repeat begin in the odd position or both begin in the even position, the ciphertext repeats. If one occurrence is in an odd position and one is in an even position, they will produce different ciphertext. As a result, nearly half of all plaintext repeats are suppressed. This is shown in these three alternate examples, all enciphered from Table 6-1.

```
   at ze ro fo ur ze ro ze ro st op
   PC CV EJ PJ DF CV EJ CV EJ DC CI

-a tz er of ou rz er oz er os to p-
-- OS UF PU RB LS UF GS UF SD TJ --

-a tz er ot hr ee ze ro ze ro st op
-- OS UF TC YF RV CV EJ CV EJ DC CI
```

c. In the first example, all three *ZEROs* produce a repeat when they all begin in the even position. In the second example, they all begin in the odd position, and only the portions of the three *ZEROS* that appear as complete digraphs (the ERs) produce a repeat. In the third example, the two *ZEROs* that begin in the even position produce repeats, but the first *ZERO,* which begins in the odd position, does not.

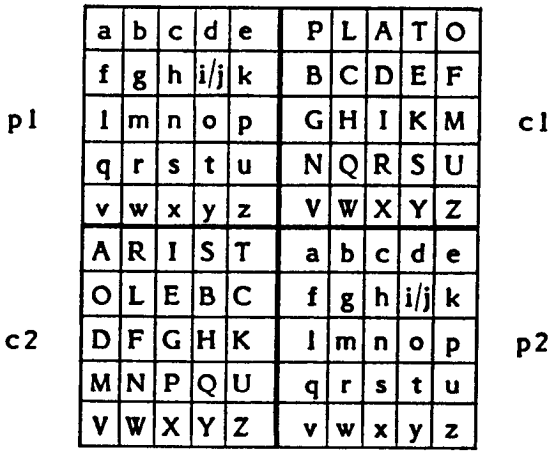
d. The suppression of individual letter frequencies and a significant portion of plaintext repeats means that digraphic systems are considerably more secure than unilateral systems and most multiliterals.

6-3. Four-Square System

Large table digraphics are awkward systems for military usage. In their place, there are several much more convenient small matrix digraphic systems available with about the same degree of security. The first of these is the four-square.

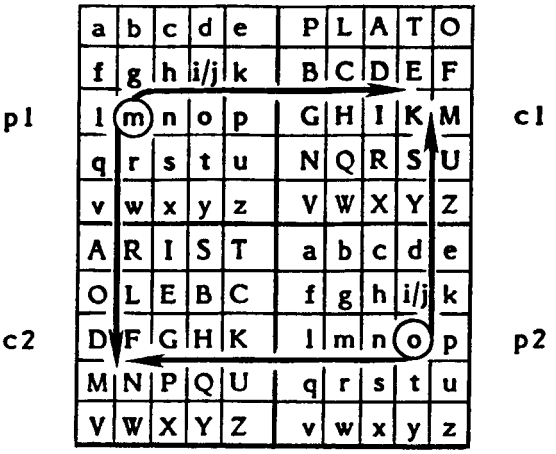
a. The four-square consists of four 5 by 5 matrices in a square. The two plaintext letters and the two ciphertext letters of each encipherment each use a different

square. The squares marked p1 and p2 usually, but not always, contain standard sequences. The two squares marked c 1 and c2 can include any mixed sequence.

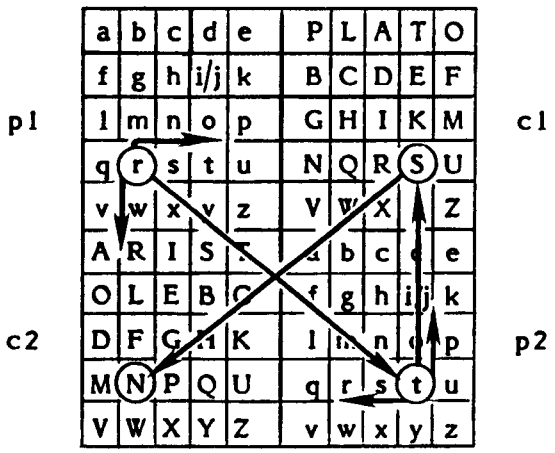


p: mo rt ar fi re
c: KF SN LM EO UR

b. Encipherment or decipherment follows a rectangular pattern. Whether enciphering or deciphering, the letters of the digraphs are located in the appropriately labeled squares. These letters form diagonally opposite corners of a rectangle. The equivalents, plaintext or ciphertext, are the remaining corners of the same rectangle. For example, plaintext MO determines the rectangle outlined in the square below. Plaintext M determines the upper row and the left column of the rectangle. Plaintext O determines the bottom row and the right column of the rectangle. The ciphertext equivalent, KF, is then found in the remaining corners in the appropriately labeled squares.



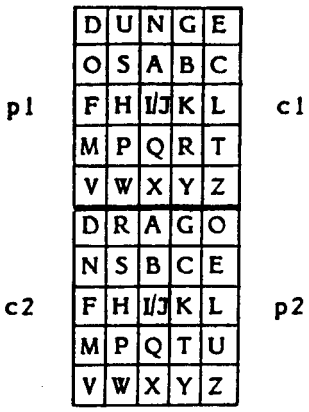
c. For a second example, to encipher RT, R is located in the pl square, and T is located in the p2 square. The ciphertext equivalent of RT is found in the remaining corners of the rectangle prescribed by RT. The first ciphertext letter, S, is found in the cl square in the plaintext T column and the plaintext R row. The second ciphertext letter, N, is found in the c2 square at the intersection of the plaintext R column and the T row. Tracing the letters from pl to p2 to cl to c2 is shown below.



d. Decipherment is handled in exactly the same way, except that the ciphertext letters in the cl and c2 squares determine the rectangle by which the plaintext letters are found.

6-4. Vertical Two-Square

The two types of two-squares are simpler than the four-square system. The first is the vertical two-square, which uses two 5 by 5 matrices one on top of the other. Normally both squares contain mixed sequences.



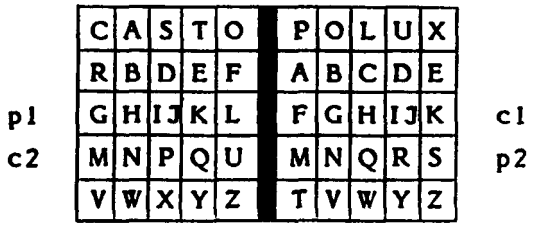
p: al lq ui et on th ew es te rn fr on tx
c: CJ IU NH GU ON PL UZ UE TE MC HD ON QZ

a. The rectangular rule used with the four-square is used with the two-square, also. Whenever the letters to be enciphered are in the same column, however, the letters become their own equivalents. The encipherment of ON and TE in the example illustrates this.

b. The case where the plaintext letters remain unchanged in the ciphertext is called a transparency. A weakness of this system is that in the long run, about 20 percent of the digraphs in a cryptogram will be transparencies. This is enough to give away more plaintext in many cases and enable a speedy solution.

6-5. Horizontal Two-Square

The second kind of two-square is the horizontal two-square, like the vertical, it uses two 5 by 5 matrices.



p: we ha ve no ty et be gu nt of ig ht
c: ZB FB ZR NA UY AY EB JC MW PL GI FW

a. The rectangular rule again applies. In the horizontal two-square, values on the same row are replaced with the same letters in the reverse order. This is illustrated by the encipherment of the plaintext letters *be* and *ig* in the example.

b. Digraphs in ciphertext which are the same as the plaintext in reverse, are called reverse transparencies. Like the direct transparencies of the vertical two-square, they occur in the long run in about 20 percent of the digraphs. They severely weaken the security of the system.

6-6. Playfair Cipher

The Playfair cipher is the most common digraphic system. *Playfair* is always capitalized, because it was named for a Lord Playfair of England. It is the simplest of systems to construct, using only a 5 by 5 matrix, yet it is more secure than uniliterals and most multiliterals. The rules of encipherment and decipherment are a little more complex than the previous digraphic systems. Sizes other than 5 by 5 are occasionally used.

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

p: th es ho th ea rd ro un dt he wo rl dx
c: QB CU PQ QB NE AJ DT ZU RO CP VQ GM GV

a. The first rule of encipherment and decipherment is the familiar rectangular rule. This applies any time the two letters to be enciphered or deciphered are not in the same row or column. The first four digraphs in the example follow this rule. One additional step must be remembered. In tracing the encipherment or decipherment in the matrix, always move vertically from the second letter to the third letter. For example, to encipher TH, locate the T and the H and move vertically from the H to the letter that is in the same column as the H and the same row as the T. Following this rule, TH is enciphered as QB, not BQ. Similarly, to decipher CU, locate the C and the U, move vertically from the U to find the first plaintext letter E and then the second plaintext letter S.

b. When the two letters to be enciphered or deciphered are in the same row, follow the rule, *encipher right, decipher left*. To encipher or decipher, pick the letter to the right or left of each letter of the given digraph, as appropriate. In the example, the plaintext letters R and D are in the same row. They are enciphered with the letters immediately to the right of each letter, producing ciphertext AJ (or AI). If a letter to be enciphered is at the right edge, as in the encipherment of HE, the next letter to the right of the right edge is considered to be the letter in the same row at the far left. The letter to the right of E is P. Similarly, if deciphering, the letter to the left of the left edge is the letter at the far right in the same row. The letter to the left of F is N. Each row is treated as if it were written in a circle with the first letter of a row immediately following the last letter.

c. When the two letters to be enciphered or deciphered are in the same column, use the rule *encipher below, decipher above*. To encipher EA in the example, the letters below E and A are N and E respectively. To decipher ZU, the letters above Z and U are U and N respectively. As with the rows, columns are treated as if they were written in a circle. The letter after the bottom letter in a column is the top letter; the letter before the top letter is the bottom letter.

d. The rules *encipher right, decipher left* and *encipher below, decipher above* produce the acronyms ERDL and EBDA. For many analysts, it is convenient to memorize these pronounceable acronyms to remember the rules.

e. The rectangular rule and the row and column rules take care of all possible cases except double letters. In the Playfair system, there is no rule for enciphering or deciphering a double letter in the same digraph. When double letters are encountered in plaintext in the same digraph, the cryptographer must break up the double letters with a null letter, such as inserting an X between them. As a result, double letters will never be encountered in the ciphertext, except in error. This is only true of the Playfair system. Four-squares and two-squares can handle double letters without any problem.

Section II
Identification of Polygraphic Substitution

6-7. General Digraphic Characteristics

Certain identifying characteristics are common to all digraphic systems. Other characteristics appear only with specific systems.

a. Message lengths, repeats, and distances between repeats are likely to be even in length in all digraphic systems because the basic unit is two-letters. Furthermore, the systems which use 5 by 5 matrices will often only use 25 letters, omitting either the I or the J in ciphertext. In some cases, these values will be used alternately just to ensure use of all letters.

b. Digraphic systems are most often mistaken for biliteral with variant systems, because both exhibit ciphertext which breaks into units of two and both can use most letters. The key distinction to look for between biliterals and digraphics is the complete absence of any positional limitation (paragraph 5-5b) in digraphic systems.

c. Two-square systems stand out because of the director reverse transparencies. Scan the text for the presence of good plaintext digraphs, either direct or reversed, to identify two-square systems. Direct transparencies indicate vertical two-squares; reversed transparencies indicate horizontal two-squares.

d. If no double letters are present in a digraphic, it is probably a Playfair system.

e. Monographic frequency counts for digraphic systems are not as flat as random text and not as rough as plaintext or unilateral systems. They generally fall in between the two. The monographic phi test can be used to confirm this, if necessary.

6-8. Digraphic Frequency Counts

There are several types of frequency counts you can take for working with digraphic systems.

a. The most common way to take a digraphic count is to break the text into digraphs and count those digraphs. For example, given text ABCDE FGHIJ . . . , you would normally break it as AB, CD, EF, GH, IJ, ... There are two other ways to take a digraphic count, however. If you are unsure whether there may be indicator groups or null letters at the beginning, you may not know where to begin breaking the text into digraphs. As a comparison, you can skip the first character and begin separating the text into digraphs beginning with the second character. This will produce a completely different set of digraphs than the usual method: A, BC, DE, FG, HI, J . . . . The third way to produce a digraphic count is to combine the two methods to count all possible digraphs. In this case, you would count AB, BC, CD, DE, EF, FG, GH, HI, IJ, ... Unless you have a reason to want an alternate method, stick to the first method.

b. There are two ways to record your count on paper. One is to make a 26 by 26 square on graph paper, and mark the digraphs in the appropriate cells. The other way, useful with short cryptograms, is to write the letters A through Z horizontally, and mark the digraphs by putting the second letter of each digraph under the first letter of the digraph in the A through Z sequence. Then by scanning the columns under each letter for repeated letters, you can readily spot repeated digraphs. This method takes much less space than a 26 by 26 square and gives you the same information.

6-9. Digraphic Coincidence Tests

The phi test and phi index of coincidence can be calculated for digraphic frequency counts as well as monographic.

a. The digraphic phi test is calculated in essentially the same way as the monographic test. In the monographic phi test, 1 out of 26 comparisons in random text was expected to be a coincidence for a probability of 0.0385. In the digraphic phi test, 1 out of 676 comparisons is expected to be a coincidence for a probability of 0.0015. The

probability of a coincidence in plaintext is 0.0069 instead of 0.0667. Thus, the formulas for the digraphic phi test are—

$$2\,\phi p = 0.0069\ N\ (N - 1).$$

$$2\,\phi r = 0.0015\ N\ (N - 1).$$

$$2\,\phi o = \Sigma f\ (f - 1).$$

$$2\,\Delta IC = \frac{676\ \Sigma f\ (f - 1)}{N\ (N - 1)} = \frac{2\,\phi o}{2\,\phi r}.$$

N is the total number of digraphs counted.
The frequency of each repeated digraph is f.

b. As discussed in the first part of this chapter, digraphic ciphertext frequencies will occur with the same numbers as plaintext frequencies when digraphic systems are used. If the digraphic ϕo is close to ϕp but the monographic ϕo is low, the system is likely to be a digraphic system. If you are using the index of coincidence form of the test, the expected $2\ \Delta IC$ is 4.6. The results are much more variable than the monographic test, because of the large number of different elements counted, but it can still be used as a guide. As with any statistical test, the results should not be used by themselves, but used along with all other available information.

6-10. Examples of System Identification

Three messages in unknown systems follow to show the process that leads to system identification. Repeats are underlined, monographic and digraphic frequency counts are shown, and monographic and digraphic ICs are calculated for each. The three messages were all sent by the same headquarters to subordinate elements, and all contained a common message serial number in their header.

a. Message texts and data.

Message 1:

TVCX XSWM WZWV JEVH HCJS IUZZ TVKP VYUY JWTZ CUIK
XCEI SVJC XIUT IDDI ETWM IWHH ISWC TIXP ZTVK RIKU
IKCU ISDV UHVM IRPC WUTU CJZK VUTV JTNI XMIB VYUZ
JVTW EIZT VKEC JEIX CCXX XICM IZEV HHCK CZZI ZEVH
HCCJ SYJJ IEIZ ZCUP HISW ECXK UVEI SYUI ZZTV KKIJ

AUII J

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	1	19	3	11	0	0	10	28	13	11	0	5	1	0	4	0	2	8	13	15	18	10	11	5	16

Total letters = 205

Monographic IC = 1.74

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	1	0	0	0	0	0	2	1	0	1	0	0	0	0	0	0	0	2	0	0	1	0	1
D	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
E	0	0	2	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H	0	0	2	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
I	0	1	0	1	1	0	0	0	1	1	2	0	0	0	0	0	0	1	2	0	1	0	1	1	0	2
J	0	0	1	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0
K	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	2	0
T	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	4	1	0	0	1
U	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	1	0	1	0	0	1	1
V	0	0	0	0	0	0	0	2	0	0	2	0	1	0	0	0	0	0	0	1	0	0	0	2	0
W	0	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	1	0	0	0	1
X	0	0	1	0	0	0	0	0	2	0	1	0	1	0	0	1	0	0	1	0	0	0	1	0	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Z	0	0	1	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	2	0	0	0	0	0	2

Total digraphs = 102

Digraphic IC = 3.41

Message 2:

NPEG MISY DQQR PATH GFTS LYUV DNPR RWIP SPDR AGYL
RKBE FIPO EGLY RFCZ AFFP SYLE KZLF SDFN LRVI NPOC
CRYL NCYL FMPT HTYA IWES TNNE VARP TNPO OZLR YAOW
IPAV PNUE AINP XKGV EFGE EGKY RLGS AIBP KZGF NCUV
IAUA THGF GVSI PVRA EFUV AGYI LFSD EBKR TPEF SIYL

UVDN PRLA VNYL ARXX

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
15	3	5	6	13	14	12	3	12	0	6	14	2	13	5	18	2	15	10	8	6	11	3	3	13	4

Total letters = 216

Monographic IC = 1.26

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	0	0	0	0	0	1	2	0	2	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
B	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
D	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	1	0	0	0	0	0	0	0	0
E	0	1	0	0	0	3	3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
I	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0	0
J	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	2
L	1	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	2	0
M	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
O	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
P	1	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	2	0	1	0	1	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
R	1	0	0	0	0	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0
S	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0
T	0	0	0	0	0	0	0	2	0	0	0	0	0	2	0	1	0	0	1	0	0	0	0	0	0	0
U	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0
V	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
X	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Y	2	0	0	0	0	0	0	0	1	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Z	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Total digraphs = 108

Digraphic IC = 5.38

Message 3:

GMGH NGMO RWOG GOEG HWMM HOHR GLNM GEGG HDND HADD
OONL MFRM GFER MLEE GEYO NANW GAGW GFRF YDYL DOMA
MRYG YFOW ODGR HLNG RWDW YAGM OOOL OAOW NFHM GOAD
DOGW GDHG DWDG HOYD GMOO OWAR MMHM GERL NEOO RANL
DWRL NDNA DOOG DLHR YLHG HEED OWYR ERNG HWYA HFYL

YGGL RFML GRYA HFHE GAGM EOOW RWAG DOOM GRNW NLMF
HLEH GFGO YMOW RMHF GERA NMYD HAYF OORW NGYD MWRO
MODW NDEG DOMM YMHR GGHD YDMA NGMF RMDW MMNF HEHD
GHND YGGL ODYW GAHL OONF OWRF MMYG YAAE HDDO DDHW
YMNG MORL YLGE YFDW DGNO NAOO MFRM HMGR RAOE DOGL

DRNL OWDO HAXX

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
23	0	0	40	20	19	54	32	0	0	0	22	40	26	50	0	0	31	0	0	0	0	27	2	26	0

Total letters = 412
Monographic IC = 2.16

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	2	0	0	2	0	0	0	0	1	0	0	8	0	0	1	0	0	0	0	6	0	0	0
E	0	0	0	1	1	0	2	1	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	3	0	0	1	5	3	2	2	0	0	0	4	4	0	3	0	0	4	0	0	0	0	2	0	0	0
H	3	0	0	4	3	3	2	0	0	0	0	3	3	0	2	0	0	3	0	0	0	0	3	0	0	0
I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	2	0	0	0	0	4	0	0	0	0	0	2	5	0	3	0	0	1	0	0	0	0	1	0	0	0
N	3	0	0	4	1	3	6	0	0	0	0	4	2	0	1	0	0	0	0	0	0	0	2	0	0	0
O	1	0	0	2	1	0	2	0	0	0	0	1	1	0	7	0	0	0	0	0	0	0	8	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
R	3	0	0	0	0	3	0	0	0	0	0	3	4	0	1	0	0	0	0	0	0	0	4	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Y	4	0	0	5	0	3	4	0	0	0	0	4	3	0	1	0	0	1	0	0	0	0	1	0	0	0
Z	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Total digraphs = 206
Digraphic IC = 8.90

b. Different analysts might approach the identification of the systems used in these messages in different ways, but here is one example of how the systems can be identified.

(1) Although the messages all carry the same message serial number, which is usually a sign of isologs, the messages are all different lengths. If they are isologs, they are not enciphered in the same system.

(2) A comparison of monographic frequency counts confirms that they are in different systems. The highs and lows in each frequency count are too different for any possibility of repeated use of the identical system.

(3) The ICs give a different picture in each. Message 1 has monographic and digraphic ICs consistent with plaintext or a unilateral system. The digraphic IC of 3.41 is slightly below the expected 4.6, but it is within acceptable limits. Message 2 shows a low monographic IC of 1.26, but the digraphic IC of 5.38 is also well within plaintext limits. This is typical of digraphic systems. Message 3 is quite high in both monographic and digraphic ICs.

(4) Messages 1 and 2 use nearly all letters. Message 3, which is twice as long as message 1, uses only 14 different letters. The high ICs and the limited letter usage are consistent with a biliteral with variants system. A close inspection of the digraphic frequency count will show rows and columns with very similar patterns, suggesting external variants that can be combined. Different letters are used in the row position than those used in the column position. This positional limitation confirms the identification of a biliteral with variants system.

(5) Message 1 has the most repeated text, which is consistent with a unilateral system. Message 2 has only a few repeats and message 3 has only short and fragmentary repeats. In message 3, the fragmented repeat on lines 7 and 10 are in the identical relative position in message 2 as the ZTVK repeat in lines 2 and 5 of message 1. This similarity strongly confirms that the two messages are isologs.

(6) The identifications of the systems in messages 1 and 3 are clear at this point, but message 2 still needs to be•clarified. The underlined repeats in message 2 are in the same relative position as in message 1, if you adjust for the slightly increased length of the message. Only some of the repeats from message 1 appear in message 2, however. This is consistent with a digraphic system, which will only show repeats that begin in the same even or odd position.

(7) In message 2, a check of the long diagonal from the AA position to the ZZ position of the digraphic frequency count shows that the only double letter that appeared was the filler XX at the end of the message. The Playfair is the only

digraphic system which will not show double letters. Finally, because the Playfair cannot encipher double letters, all double letters that occur in digraphs must be broken up by the insertion of null letters. This characteristic explains how it can be an isolog, but appear slightly longer. The three messages are all clearly isologs, and the systems are confidently identified, lacking only the final solution for full confirmation. Solution techniques for each of the major digraphic system types are explained in the next chapter.