Mathematics 124 (Winter 2009)

Cryptanalysis of Monoalphabetic Substitutions

Note: There are a few discussions in the textbook about the cryptanalysis of a monoalphabetic substitutions. In particular, read pages 57–59; Example 2.2.7 on pages 77–80; Example 2.3.3 on pages 85–90.

Example: Consider the following ciphertext produced by a monoalphabetic substitution:

```
RADHA JCRWC MJCOA NZCSY JVHCY MVAGH WZHSM
LTCOH WCAVA SPDLO JLGHV ZASPV LZCII HSLOZ
                 CVNZ
```

Assumptions, Information, Strategy

- It is in English, and has roughly the same *statistics* as standard English.
- Word divisions are not preserved.
- Use frequency analysis to guess high frequency letters: E, T, N, O, R, I, A, S make up 70% of letters in English.
- Try to identify vowels.
- Try to identify digraphs.
- Use cribs.
- Guess! and rely on luck.

Table 2.6: Relative Letter Frequencies in a Sample of English

letter	frequency (%)	letter	frequency (%)
A	8.399	N	6.778
B	1.442	O	7.493
C	2.527	P	1.991
D	4.800	Q	0.077
E	12.150	R	6.063
F	2.132	S	6.319
G	2.323	T	8.999
H	6.025	U	2.783
I	6.485	V	0.996
J	0.102	W	2.464
K	0.689	X	0.204
L	4.008	Y	2.157
M	2.566	Z	0.025

Table 2.7: Commonest Digraphs and their Frequencies in a Sample of English

digraph	frequency (%)	digraph	frequency (%)
TH	3.319	ES	1.213
HE	2.859	TO	1.213
IN	2.081	NT	1.200
ER	1.596	EA	1.059
ED	1.493	OU	1.047
AN	1.430	NG	1.034
ND	1.430	ST	1.034
AR	1.302	AS	0.9957
RE	1.302	RO	0.9957
EN	1.289	AT	0.9829

(Part of) Table 2.8: Commonest Trigraphs and their Frequencies in a Sample of English

trigraph	frequency (%)	trigraph	frequency (%)
THE	1.82	ING	0.68
AND	0.77	HER	0.50

When we analyze the ciphertext, we see that the commonest letters, digraphs, and trigraphs are:

C	A	H	V	Z	L	S				
9	7	7	6	6	5	5				
ZC	AS	CO	GH	HW	HS	JC	LO	NZ	SP	VA
3	2	2	2	2	2	2	2	2	2	2
ASP										
2										

That is, we guess

plaintext ↦ ciphertext

{E, T, N, O, R, I, A, S} ↦ { C, A, H, V, Z, L, S, T, O}

{TH, HE, IN, ER, RE, ON, AN, EN, AT} ↦ {ZC, AS, CO, GH, HW, HS, JC, LO, NZ, SP, VA}

{THE, AND, ING, HER} ↦ {ASP}

Some Known Cribs: F ↦ D and H ↦ Z

And here is the ciphertext with the word divisions put back in:

```
RADH AJ CRWCMJ C OANZ CSY JVHCYM
VAGH WZHS MLT COH WCAVASP DLO
JLGHVZASP VL ZCIIHS LO ZCVNZ
```

The quote is from Charlotte's Web by E.B. White.

Note: This is an example of a **cryptogram**, a short piece of text encrypted with a simple monoalphabetic substitution cipher. To solve the puzzle, one must recover the original lettering. Though once used in more serious applications, they are now mainly printed for entertainment in newspapers and magazines. A useful tool for solving cryptograms may be found at

http://scottbryce.com/cryptograms/index.htm