Cryptography Take 1

Blake Thornton and Louis Beaugris

1 Substitution Ciphers

One of the simplest ways to encode text is to use a substitution cipher. Here we substitute one letter for another. Often this is arranged by shifting the alphabet. A straight shift is called a Caesar cipher. Here is a Caesar cipher with a shift by 3:

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

We can encode using this substitution. A $\mapsto$ D, B $\mapsto$ E, etc.

Message:	I		L	O	V	E		M	A	T	H
Code:	L		O	R	Y	H		P	D	W	K

1. The following was encoded with a Caesar cipher and shift 3. Decode it:

   FDHVDU FLSKHUV DUH QRW VHFXUH

2. Why did the mathematical tree fall over? (Shift of 7)

   ILJHBZL PA OHK UV YLHS YVVAZ

3. What does a bull add with? (Shift of 12)

Message:												
Code:	M		O	A	I	O	G	X	M	F	A	D

4. What did one math book say to the other? (Shift of 17)

   UFE'K SFKYVI DV Z'MV XFK DP FNE GIFSCVDJ!

2 Breaking a Caesar Ciphers

If you know a Caesar cipher was used, then you can just try all possible shifts until the message starts making sense.

5. If you use the set $\{A, B, C, \ldots, Z\}$, how many different Caesar ciphers are there?

   In other words, if you want to break a Caesar cipher, how many tries might it take if you are really unlucky and try the correct shift last?

Try to break the following codes that were coded with a Caesar cipher.

6. What goes up but not down?

Message:								
Code:	C	S	Y	V		E	K	I

7. ESP DXLWW HZCOD LCP GPCJ SPWAQFW

8. GUVFVFUNEQREORPNHFRGURERNERABFCNPRF

3 Letters to Numbers

Ideally, we would like to use some mathematics in our coding. To do this, first, we need to translate our letters to numbers. Lets agree to do it this way:

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
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

Lets encode "A BLUE AND RED COW" using a Caesar cipher with a shift of 12. We just add 12 to the numbers. The only problem is when the numbers are over 25.

Message:	A		B	L	U	E		A	N	D		R	E	D		C	O	W
Numbers:	0		1	11	20	4		0	13	3		17	4	3		2	14	22
Add 12	12		13	23	6	16		12	25	15		3	16	15		14	0	8
Code:	M		N	X	G	Q		M	Z	P		D	Q	P		O	A	I

Once you get your text into numbers, there is a lot more you can do. For example, you could encode your message with something like $x \mapsto 3x + 7$.

9. Take the code A BLUE AND RED COW and encode it using various formulas. We did the $3x + 7$ code for you, you do the others.

Message:	A		B	L	U	E		A	N	D		R	E	D		C	O	W
Numbers:	0		1	11	20	4		0	13	3		17	4	3		2	14	22
$3x + 7$	7		10	14	15	19		7	20	16		6	19	16		13	23	21
$3x + 7$ Code	H		K	O	P	T		H	U	Q		G	T	Q		N	X	V
$2x + 1$																		
$2x + 1$ Code:																		
$7x + 13$																		
$7x + 13$ Code:																		

10. One problem is decoding. For the $3x + 7$ code, you can decode by using the formula $9x - 11$. Try it. Here is a message that we encoded with the $3x + 7$ code. Decode it with $9x - 11$.

Code:	M	C	P	H	Q	T		J	X	Q	V	F	P
Code Numbers:													
$9x - 11$:													
Decoded Message:													

11. Can you find the decoding formula for the coding formula of $9x + 2$? Use your formula to decode the answer to this question.

    Which word in the dictionary is spelled incorrectly?

Code:	W	P	U	Y	Z	Z	M	U	R	X	K
Numbers:											
Decoded Numbers:											
Message:											

12. Look closely at the line for the $2x + 1$ code in Problem 9. There is a big problem with this line, find it.

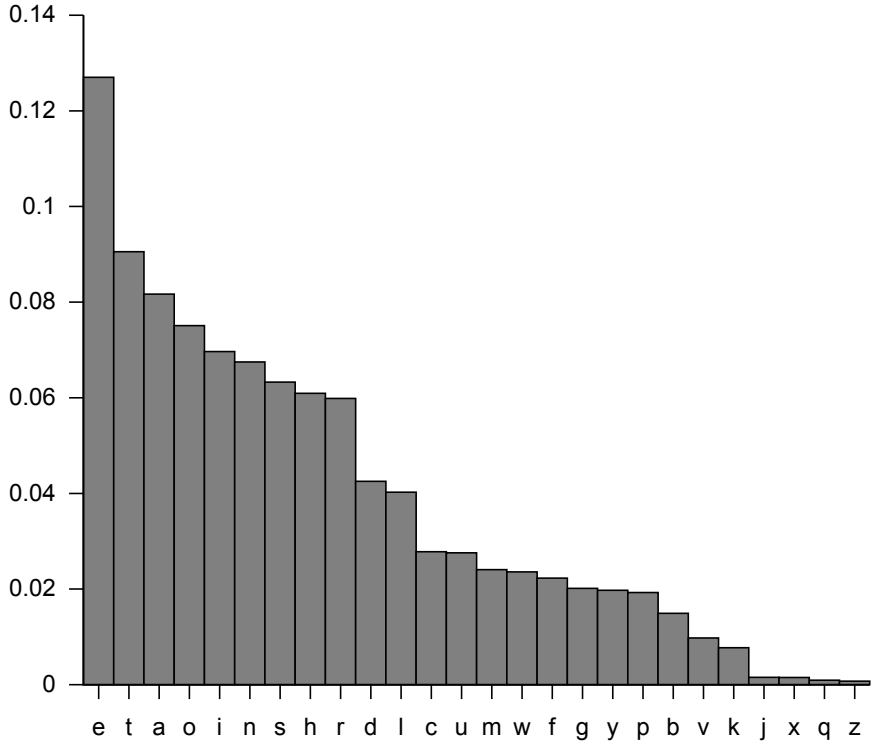
    Why did this happen? How can we avoid this problem in the future?

13. If you use the set $\{A, B, C, \ldots, Z\}$, how many different substitution ciphers are there?

4 Frequency Analysis

Substitution ciphers can be broken by using frequency analysis. Here are the rough frequencies that you expect to see in English text, ordered by their frequencies in text.

Figure 1: Frequency of Letters



To decode a text, figure out which symbol is the most common in your coded text. Since "E" is the most common letter in English, that is probably the letter to substitute for your most common symbol. The tricky part is that the frequencies that you see in actual text are often different than the ideal case. So, you might have to experiment a bit, which takes time.
The longer the text you are trying to decode, the more accurate this method is.

14. Here are a bunch of coded messages that you intercepted. You believe it is a substitution cipher, but you are not sure what the substitution is.

    Use frequency analsysis to decode. Use any other clues you can discover from the message.

- XCF MR YTLDQH CNGD HYQTIDH? HR YCDF MRAY LDY HIRYYDM.
- XCNY MTM YCD GRSVNAR HNF YR CTH XTUD? T SNGN FRP HR JPVC!
- XCNY MTM RAD XNSS HNF YR YCD RYCDQ XNSS? T'SS JDDY FRP NY YCD VRQADQ.
- XCNY MR FRP VNSS YXR MTARHNPQH YCNY CNGD EDDA TA NA NVVTMDAY? YFQNAARHNPQPH XQDVBH.
- XCDQD HCRPSM N UNN IRPAM NSTDA LR? RA N MTDY
- XCF MTM YCD ITVYPQD LR YR KNTS? EDVNPHD TY XNH UQNJDM.
- XCNY MTM YCD INIDQ HNF YR YCD IDAVTS? XQTYD RA!
- XCNY LDYH XDYYDQ YCD JRQD TY MQTDH? N YRXDS.
- XCF MR MQNLRAH HSDDI MPQTAL YCD MNF? HR YCDF VNA UTLCY BATLCYH!
- XCNY MTM YCD HYNJI HNF YR YCD DAGDSRID? HYTVB XTYC JD NAM XD XTSS LR ISNVDH!
- XCNY MTM RAD DSDGNYRQ HNF YR YCD RYCDQ DSDGNYRQ? T YCTAB T'J VRJTAL MRXA XTYC HRJDYCTAL!
- XCF XNH YCD EDSY NQQDHYDM? EDVNPHD TY CDSM PI HRJD INAYH!
- XCTVC CNAM TH TY EDYYDQ YR XQTYD XTYC? ADTYCDQ, TY'H EDHY YR XQTYD XTYC N IDA!
- XCF VNA'Y FRPQ ARHD ED EV TAVCDH SRAL? EDVNPHD YCDA TY XRPSM ED N URRY!
- XCNY CNH URPQ XCDDSH NAM USTDH? N LNQENLD YQPVB!
- XCF MTM YCD QREEDQ YNBD N ENYC EDURQD CD HYRSD UQRJ YCD ENAB? CD XNAYDM YR JNBD N VSDNA LDY NXNF!

Here are the letters counted in the text.

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
40	8	59	101	16	20	7	45	14	12	1	16	37	71	0	19	36	58	30	50	9	22	0	39	91	0

5 More Coded Message (Mostly Math Jokes) to Decode (This part is extra credit 1 pt each)

15. A Caesar cipher (you must figure out the shift).

    PAR PTL MAX FTMA UHHD LTW? UXVTNLX BM ATL LH FTGR IKHUEXFL.

16. A Caesar cipher (you must figure out the shift).

    EPIB AWZB WN TQVOMZQM LWMA I UMZUIQL EMIZ? IV ITOIM-JZI.

17. Encoded with $15x + 12$.

    ENML FO IOA YUL CJ IOA FCPCFU LNU QCHQAKJUHUZQU OJ MZ CYVOO BI CLW FCMKULUH? UWGCKO DC.

18. Encoded with $19x + 11$.

    NOLI QR ZRB VJI HC ZRB QHUHQJ IOJ XHWXBFCJWJYXJ RC LY HVMRR EZ HIP QHLFJIJW? JPTHFR KH.

19. Encoded with $25x - 4$.

    APY TOT DPS VIY SWD POE KWDP PIKSAIFM? VSUWCES DPS DSWUPSF DILT POK OD AWE W HOSUS IR UWMS.

20. Encoded with $5x$.

    GJAR OM A IARJ RUAKJUH'M ZABSHORU MWI? MWIIUH!

21. Encoded with $5x + 10$.

    BTERE KRE PK BAHEW CJ SKBTESKBYUYKXW. BTCWE QTC IXCQ PYXKRA KXZ BTCWE QTC ZCX'B.

22. Encoded with $21x + 20$.

    U OCMUH LUI IATAH FUYQLDANI, UHF AUKL FUYQLDAN LUI U PNCDLAN. LCO MUHE KLGRFNAH FCAI DLA OCMUH LUTA URR DCQADLAN? AGQLD.

23. Encoded with $7x + 2$.

    AZCF AWMBX OWM YCO GL YWIEWPE'Y DCRRWF ZCX XGEX? DWBOSWP.