CLASSICAL CRYPTOGRAPHY COURSE
BY LANAKI
October 23, 1995
Revision. 0

LECTURE 2
SUBSTITUTION WITH VARIANTS
Part I

SUMMARY

In Lecture 2, we expand our purview of substitution ciphers, drop the requirement for word divisions, solve a lengthy Patristocrat, add more tools for cryptanalysis, look at some historical variations and solve the assigned homework problems.

IDENTIFYING SUBSTITUTION AND TRANSPOSITION CIPHERS

Recall from Lecture 1, that the fundamental difference between substitution and transposition ciphers is that in the former, the normal or conventional values of the letters of the PT are changed, without any change in the relative positions of the letters in their original sequences, whereas in the latter, only the relative positions of the letters of the PT in the original sequences are changed, without any changes to the conventional values for the letters.

I used the term uniliteral frequency distribution [UFD] (I also misspelled uniliteral as unilateral) to identify the simple substitution cipher. Three properties can be discerned from the UFD applied to CT of average length composed of letters: (1) Whether the cipher belongs to the substitution or transposition class; (2) If to the former, whether it is monoalphabetic or non-monoalphabetic in character, (3) If monoalphabetic, whether the cipher alphabetic is standard (direct or reversed) or mixed.

CIPHER CLASS

Because a transposition cipher rearranges the PT, without changing the identities of the PT, the corresponding number of vowels (A,E,I,O,U,Y), high frequency consonants (D,N,R,S,T), medium-frequency consonants (B,C,F,G,H,L,M,P,V,W) and especially, low-frequency consonants (J,Q,X,Y,Z) are exactly the same in the CT as they are in the PT. In a substitution cipher, the conventional percentage of vowels and consonants in the CT have been altered. As messages decrease in length there is a greater probability of departure from the normal proportion of vowels and consonants. As messages increase in length, there is lesser and lesser departure from normal proportions. At 1000 letters or more, there is practically no difference at all between actual and theoretical proportions. Friedman presents charts showing the normal expectation of vowels and high, medium, low and blanks for messages of various lengths. For example, for a message of 100 letters in plain English, there should be between 33 and 47 vowels (A,E,I,O,U,Y). Likewise, there will be between 28 and 42 high-frequency consonants (D,N,R,S,T); between 17 and 31 medium frequency consonants (B,C,F,G,H,L,M,P,V,W); between 0 and 3 low-frequency consonants (J,Q,X,Y,Z); and between 1 and 6 blanks theoretically expected in distribution of the PT. Cipher class is considered transposition if the above limits bound the CT message and substitution if the above expected limits are outside the chart limits for the message length in question. [FR1/ p32-39]

UFD

The uniliteral frequency distribution (UFD) may be used to indicate monoalphabeticity. The normal distribution shows marked crests and troughs by virtue of two circumstances. Elementary sounds which the symbols represent are used with greater frequency. This is one of the striking characteristics of every alphabetic language. With few exceptions, each sound is represented by a unique symbol. The one-to-one mapping correspondence between PT and CT will dictate a shifted UFD with different absolute positions of the crests and troughs from normal. A marked crest-and-trough appearance in the UFD for a given cryptogram indicates that a single cipher alphabet is involved and constitutes one of the tests for a mono-alphabetic substitution cipher.

The absence of marked crests and troughs in the UFD indicates that a complex form of substitution is involved. The flattened out appearance of the distribution is one of the criteria for rejection of a hypothesis of monoalphabetic substitution.

LAMBDA BLANK EXPECTATION TEST - LB^

Friedman presents a chart supporting the LB^ test for blanks in English messages up to 200 letters. [FR1] Soloman Kullback derives the Lambda test and presents extensive probability data on English, French, German, Italian, Japanese, Portuguese, Russian and Spanish. [KULL] Statistical studies show that the number of blanks in a normal PT message is predictable. Friedman's chart shows that the plaintext limit, P and the random expectation, R limits are a function of message size. On his chart, random assortment of letters correspond to polyalphabetic CT. The number of alphabets used is large enough to approximate a UFD identical to a distribution of letters picked randomly out of a hat.

PHI TEST FOR MONOALPHABETICITY

This test compares the observed value PHI(o) for the distribution being tested with the expected value PHI(r) random and the expected value of PHI(p) plain text. For English military text,

PHI(r) = .0385N(N-1)
PHI(p) = .0667N(N-1)

where N is the number of elements in the distribution. The constant .0385 is 1/26 decimal equivalent and constant .0667 is the sum of squares of the probabilities of occurrence of the individual letters in English PT. [FR3]

Example 1 of the PHI test on the following cryptogram is:

```
O W Q W Z   A E D T D   Q H H O B   A W F T Z   W O D E Q
T U W R Q   B D Q R O   X H Q D A   G T B D H   P Z R D K

f:       3 3   7 2 1 1 4     1       4 1 6 3   4 1   5 1   3
CT:      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
f(f-1): 6 6  42 2 0 0 12    0      12 030 6  12 0  20 0   6
```

N = number of letters = sum fi = 50

PHI(o) = sum [fi(fi-1)] = 154

PHI(r) = .0358N(N-1) = .0385x 50 x49 =94
PHI(p) = .0667N(N-1) = .0667x 50 x49 =163

Since PHI(o), 154, more closely approximates PHI(p) than does PHI(r), we have mathematical corroboration of the hypothesis that the CT is monoalphabetic.

Example 2: Given the frequency distribution of CT as:

```
f:       1     1 2 3 4 2     1       4 2   1           1   3
CT:      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
f(f-1): 0     0 2 612 2     0      12 2   0           0   6
```

N = 25 letters
PHI(o) = 42

PHI(r) = 0.0385x25x24 = 23
PHI(p) = 0.0667x25x24 = 40

Since PHI(o) observed is closer to PHI(p), then this letter distribution is monoalphabetic. But compare to example 3 with 25 letters:

```
f:          1       1 1 2 1 1 1 3 1 1   1 2   1 1   1 1 2 3
CT:      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
f(f-1):     0       0 0 2 0 0 0 6 0 0   0 2   0 0   0 0 2 6
```

N = 25 letters
PHI(r) = 0.0385x25x24 = 23
PHI(p) = 0.0667x25x24 = 40

Since PHI(o) observed is closer to PHI(r), then this letter distribution is non-monoalphabetic.

Before we think this test is perfect, the student should try the above PHI test on the phrase:

" a quick brown fox jumps over the lazy dog"

He will find that N=33, PHI(o)= 20 and PHI(r) = 41; PHI(p)=70.

since the observed value is less than half of PHI random, this would suggest that the letters of this phrase could not be plain text in any language. Think about the cause of this result. For a simplified derivation, see Sinkov [SINK]

Kullback gives the following tables for Monoalphabetic and Digraphic texts for eight languages:

	Monoalphabetic Text	Digraphic Text
English	0.0661N(N-1)	0.0069N(N-1)
French	0.0778N(N-1)	0.0093N(N-1)
German	0.0762N(N-1)	0.0112N(N-1)
Italian	0.0738N(N-1)	0.0081N(N-1)
Japanese	0.0819N(N-1)	0.0116N(N-1)
Portuguese	0.0791N(N-1)	
Russian	0.0529N(N-1)	0.0058N(N-1)
Spanish	0.0775N(N-1)	0.0093N(N-1)

Random Text

Monographic	Digraphic	Trigraphic
.038N(N-1)	.0015N(N-1)	.000057N(N-1)

Note that the English plain text value is slightly less than Friedman's. [KULL] [SINK]

INDEX OF COINCIDENCE (I.C.)

Friedman made famous the Index of Coincidence. It is another method of expressing the monoalphabeticity of a cryptogram. We compare the theoretical I.C. with the actual I.C. I.C. is defined as the ratio of PHI(o)/PHI(r). Thus, in example one the I.C. is 154/94 = 1.64. The theoretical I.C. for English is 1.73 or (.0667/.0385). The I.C. of random text is 1.00 or (.0385/.0385). Friedman wrote a paper entitled "The Index of Coincidence and Its Application in Cryptography", which is perhaps the most ground breaking treatise in the history of cryptography. [FR22]

CIPHER ALPHABETS - STANDARD OR MIXED

Assuming a UFD that is monoalphabetic in character, we observe the crests and troughs of the distribution. If they occupy relative offset positions to the normal UFD, than the alphabet is most likely standard, (A, B, C,..). If not, the CT is prepared using a mixed alphabet. The direction the crests and troughs progress left to right or right to left tell us whether the alphabet is standard or reversed in direction.

LONG WORD RISTIES - SHERLAC METHOD

When an Aristocrat consists of all long words, it may be attacked by the SHERLAC Method. The object is to compare vowel positions and word endings in a columnar display of the CT by individual word. We mark all low frequency (f <= 3) , then the 2nd column position (vowel favorite) and word endings are examined. For example, from S-TUCK: [TUCK], [B201]

```
fi  14 13 12 12 10 10 8 5 5 4 4 4 3 3 3 3 3 3 2 2 2 2
CT   D  Q  I  N  O  P A L X E R V C F H M S Y J K W Z
```

F= 127 letters = sum fi

The CT presented in columnar form and marked for low frequency letters is:

```
      c  .  c  v  .  .  v  c  c  v
1.    X  W  V  I  M  S  O  Q  P  N  V
      s     c  o     h  a  n  t  i

      c  v  .  c  c  .  v  .  v  c
2.    Q  I  F  E  D  Y  I  H  O  Q,
      n  o  b  l  e  w  o  m  a  n

      .  v  .  c  v  .  v  c  c
3.    Z  I  Y  P  I  Y  N  Q  L
         o  w  t  o  w  i  n  g

      v  .  v  c  c  v  c  v  c  v  c  c  .
4.    D  K  O  L  L  D  A  O  P  D  R  E  W ,
      e  x  a  g  g  e  r  a  t  e  d  l  y

      c  v  c  .  c  v  v     v  c
5.    R  N  X  M  E  D  O  X  D  R
      d  i  s     l  e  a     e  d

      c  v  .  v  c  v  v  c  c
6.    X  I  C  D  A  D  N  L  Q .
      s  o     e  r  e  i  g  n

      .  c  v  .  v  v  v  c  v  v  c
7.    M  A  I  C  I  V  O  P  N  I  Q
         r  o     o  e  a  t  i  o  n

      v  c  v  c  c  v  c  v  c  v
8.    N  Q  I  A  R  N  Q  O  P  D ,
      i  n  o  r  d  i  n  a  t  e
      .  v  c  v  c  .  .  v  c  c
9.    F  O  Q  N  X  S  H  D  Q  P
         a  n  i  s  h  m  e  n  t
```

```
      v  c  c  v  c  c  c  v  .  v  c  c  v  .  c  v
10.   N  Q  V  I  Q  P  A  I  C  D  A  P  N  F  E  D.
      i  n  c  o  n  t  r  o  v  e  r  t  i  b  l  e

      v  .  c  v  c  .  v  c  .
11.   O  J  P  D  A  H  O  P  S,
      a  f  t  e  r  m  a  t  h

      .  v  c  c  .  v  .  v  c
12.   Z  N  Q  L -J  N  K  D  A  !!!
         i  n  g  f  i  x  e  r
```

We mark the cipher as we put forth the following thoughts. Analysis of Column 2 in the above CT, shows that I appears three times with 11 low frequency contacts. It is probably a vowel but not "i." N appears twice with 5 low frequency contacts and also in third end position 3 times. Probable vowel, may be i as in ion. Q appears twice; no low frequency contacts, follows probable vowels 7 times. Might be n. i and n are placed in the CT. Word 7 yields I = o. Word 3 yields the L = g. [Word 6 may not fit though.] Word 7 also suggest that P =t for tion. P precedes N and I 4 times, and follows O 3 times. D begins one word, ends 2, has high frequency, and is scattered. Let D = e. O contacts 6 low frequency, and precedes n 3 times, t 4 times and the t is followed by e twice. We have the 'ate' trigram, so O=a. Note that A reverses with D and contacts vowels 10 times. A=r.

Word 10 shows incontrovertible. Playing it thru with V=c,F=b, and E=l, word 2 becomes noblewoman, Y=w,h=m. Word 11 is aftermath giving two more PT equivalents of J=f, and S=h. Word 4 gives us the K=x, R=D, W=y. Word 6 is sovereign and yields the PT s. The balance of the CT can be found by check off and testing.

The message reads: Sycophantic noblewoman, kowtowing exaggeratedly, displeased sovereign. Provocation inordinate, banishment incontrovertible. Aftermath king-fixer!

I have experimented with the SHERLAC method. Even when the CT includes small words <= 4 letters, it seems to yield valuable data. Just line the CT (words 5 or more letters) in columns and ignore the shorter words. Then work back and forth with the shorter words for confirmations.

PATRISTOCRATS

When we remove the crutch of word divisions in the Aristocrat and present in standard telegraphic five letter groups, we have the "Patristocrat" or "undivided." S-TUCK gives a solution procedure for "undivideds" as well as Friedman. ELCY also discusses the Aristocrat without word divisions in her chapter 11. [TUCK], [FR1], [ELCY] Friedman's presentation is excellent and is summarized for the reader.

Given P-1:

```
SFDZF  IOGHL  PZFGZ  DYSPF  HBZDS  GVHTF  UPLVD  FGYVJ  VFVHT


GADZZ  AITYD  ZYFZJ  ZTGPT  VTZBD  VFHTZ  DFXSB  GIDZY  VTXOI
        ---                    -------------

YVTEF  VMGZZ  THLLV  XZDFM  HTZAI  TYDZY  BDVFH  TZDFK  ZDZZJ
                     -------------------------------

SXISG  ZYGAV  FSLGZ  DTHHT  CDZRS  VTYZD  OZFFH  TZAIT  YDZYG
                                               --------------
```

```
AVDGZ  ZTKHI  TYZYS  DZGHU  ZFZTG  UPGDI  XWGHX  ASRUZ  DFUID
          ----                                      -----


EGHTV  EAGXX
```

There are two basic attacks on the Patristocrat. The first method creates a triliteral frequency table and the second uses the "probable word" as a wedge into the cryptogram. The first attack follows many of the vowel - consonant splitting steps that we have looked at previously.

METHOD A: Vowel - Consonant Splitting

Step 1: Inspect/mark for long repetitions, many letters of normally low frequency, such as F, G, V, X, Z; and vowels and high frequency consonants N and R are relatively scarce.

Step 2: Prepare UFD and apply PHI tests.

```
 8 4 1 23 3 19 19 15 10 3 2 5 2 0 3 5 0 2 10 22 5 16 1 8 14 35
 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
```

PHI(p) = 3668 PHI(r) = 2117 PHI(o) = 3862 ft = 235

The marked crests and troughs and the PHI test support the monoalphabetic hypothesis. Friedman advises that "the beginner must repress the natural tendency to place too much confidence in the generalized principles of frequency and to rely too much upon them. [i.e. setting Z=e, D=t] It is far better to into effective use certain other data concerning normal plain text, such as digraphic and trigraphic frequencies."

Step 3: Prepare a special worksheet; mark reversible digraphs and trigraphs, inscribe the frequencies of the first and last 10 letters, because these positions often lend themselves more readily to attack, and note positions of low frequency CT letters.

Step 4: Prepare a Triliteral Frequency Distribution (TFD) showing One Prefix and One Suffix Letter. Examine the TFD for digraphs and trigraphs occurring two or more times in the cryptogram. Note repeated digraphs and trigraphs. For the above CT,

```
           DZ  = 9x,  DF  = 5x, DV  = 2x
           ZDF = 4x,  YDZ = 3x, BDV = 2x
```

```
     Condensed Table Of Repetitions For P-1.

    Digraphs         Trigraphs          Polygraphs

DZ - 9  TZ - 5      DZY - 4            HTZAITYDZY - 2
ZD - 9  TY - 5      HTZ - 4            BDVFHTZDF - 2
HT - 8  FH - 4      ITY - 4            ZAITYDZY - 3
ZY - 6  GH - 4      ZDF - 4            FHTZ - 3
DF - 5  IT - 4      AIT - 3
GZ - 5  VF - 4      FHT - 3
        VT - 4      TYD - 3
        ZF - 4      YDZ - 3
        ZT - 4      ZAI - 3
        ZZ - 4
```

```
          IE                                                                
          ZF                                                          HV    
          GI                                                          ZG    
          SZ                                                          IY    
          VG    DU AX                                                 ZK    
          YZ    ZZ EH                                                 IY    
          ZO    FH WH                                                 HZ    
          CZ    ZF PD GT                                              VY    
          ZT    VS TU GX                                              HC    
          ZZ    DK ZH GU                                              DH    
          ZF    VH DZ KI                                              HZ    
          BV    DM YA FT                                              IY    
          YZ    EV LZ FT                                              HZ    
          ZF    DX YA HT UD                                     AR ZH       
          IZ    VH SZ TH DX                                     YD VE       
EG        ZF    YZ MZ FT HT                                     RV VX       
XS        BV    VV BI MT AT                                     FL HZ       
GV        YZ    DG TP TL XS                                     IG VZ       
ZI        AZ    TU TA FT AT        SG          UG               JX PV       
GV YD     VF    PH FY VT OY        LV          GT               XB ZG       
ZI SG     ZS VA ZG SV VT GD ZS     HL        DZ UL              DG IY       
ZI ZD     ZY DG ZI FZ FB AT ZZ TH PV FH     XI SF     SU YP HG              
GD HZ TD FZ TF SD OH GL FO VV FZ HP VG     IG LZ     ZS -F HF              
 A  B  C  D  E  F  G  H  I  J  K  L  M  N  O  P  Q  R  S  T                
 8  4  1 23  3 19 19 15 10  3  2  5  2  0  3  5  0  2 10 22                

                   KD                                                     
                   TD                                                     
                   DY                                                     
   TE              TA UD                                                  
   AD              XD FT                                                  
   ST        ZS    ZT UF                                                  
   AF        TZ    GZ DG                                                  
   DF        ZG    DY YY                                                  
   LX        TD    TD ZT                                                  
   FM        TZ    TB GZ                                                  
   YT        ZG    JT DY                                                  
   YT    X- ZB    FJ TA                                                  
   DF    GX TD    DY OF                                                  
   TT    HA IV    ZA YD                                                  
FI FH    IW ZV    DZ DR                                                  
RZ JF    SI ZF    BD GD                                                  
GP YJ    VZ TD    GD GY                                                  
HZ LD    TO GV    PF ZJ                                                  
FP GH XG FS DS    DF DZ                                                  
 U  V  W  X  Y     Z                                                     
```

Step 5: Classify the cipher letters into vowels and consonants. As we did in the Aristocrats, again we separate high frequency letters into probable vowels and consonants. If we find A, E, I, O, and N, R, S, T, we have values for 2/3 of the cipher text letters that normally (most likely) occur in the cryptogram.

> Friedman's Table 7-B in Appendix 2 confirms that vowel combine differently from consonants. The top 18 digraphs compose about 25 per cent of English text. The letter E enters into 9 of the 18 digraphs. [FRE1]
>
> ED EN ER ES NE RE SE TE VE

The remaining 9 digraphs are:

> AN ND OR ST IN NT TH ON TO

None of the 18 digraphs is a combination of vowels. So E combines with consonants more readily than with other vowels or even itself. So if the letters of the highest frequency are listed with the assume CT =e , those that show a high affinity are likely N R S T and those that do not show any affinity are likely A I O U. In P-1., Let Z = e because it is high frequency and combines with several other high frequency letters, D, F, G. The nine next highest frequency letters and their combinatorial affinity with Z are:

Z as prefix	8	4	4	1	0
	D(23)	T (22)	F(19)	G(19)	V(16)
Z as suffix	9	5	2	5	0

Z as prefix	0	6	0	0
	H(15)	Y(14)	S(10	I(10)
Z as suffix	0	2	0	0

Step 6: Analysis of Data

CT D occurs 23 times, 18 times combined with Z, 9 times as areversal ZD, DZ. T shows 9 combinations Z, 4 in ZT and 5 in TZ. D and T must be consonants. Similarly, F, G, Y are guessed as consonants. An initial cut is:

Vowels	Consonants
Z=e, V, H, S, I	D, T, F, G, Y

Friedman's Table 6 in Appendix 2 gives us 10 most frequently occurring diphthongs: [FRE1]

Diphthong:	io	ou	ea	ei	ai	ie	au	eo	ay	ue
Frequency:	41	37	35	27	17	13	13	12	12	11

Also, O is usually the vowel of second highest CT frequency. Looking at V, H, S, I not = i, can we find the CT equilvalent of PT o?

List the combinations of V, H, S, I and Z=e in the message. We examine the combinations they make among themselves and with Z = e.

ZZ = 4 VH = 4 HH = 1 HI = 1 IS = 1 SV = 1

Now, ZZ = ee. HH is oo, because aa, ii, uu are practically non-existent. oo is the second highest frequency double vowel next to ee. If H=o, then V =i, where VH occurs twice and io is a high frequency diphthong in English. So our analysis results (unconfirmed) so far are:

Z = e, H = o, V = i

So I and S should be a and u. Here we use another Friedman tool to look at the possibilities. We define the alternative PT diphthongs and add frequency values as a set.

```
  1) either I = a and  S = u,      each digraph occurs 1x
  2) or     I = u and  S = a.

HI = oa   value =  7          HI = ou  value = 37
SV = ui   value =  5          SV = ai  value = 17
IS = au   value = 13          IS = ua  value =  5
                ====                         ====
Total             25                           59
```

Alternative two seems more likely. A more precise method for choosing between alternative groups of Digraphs by considering logarithmic weights of their assigned probabilities, rather than PT frequency values. These weights are given by Friedman in [FRE2] Appendix 2. The method is detailed on pp 259-260. Tables 8 and 9A - C give the data for 428 digraphs based on 50,000 words of text. See also [KULL].

```
HI = oa   L224  = .48         HI = ou  L224  = .79
SV = ui   values= .42         SV = ai  values= .64
IS = au         = .59         IS = ua        = .42
                =====                        =====
Total Log base   1.49                         1.85
           224
```

Multiple occurrences of a digraph would be multiplied by its log base 224 relative weight and added as a group.

So we now have Z = e , H = o, V =i, S = A, I = u for vowel equivalents.

The consonants may be viewed from their combination with suspected vowels. Since VH = io might infer sion or tion tetragraphs we look at the CT and find

GVHT and FVHT

T most likely is the n and G or F could be s or t. Note that the CT D is neither PT t or n or PT s. The reversal with Z =e, suggests the letter r.

As an alternative, the Consonant-Line approach would yield

```
   B C E J K M O R W
----------------------
        Y É
      D D ÉD D D        vowel ?
      S S ÉS S         vowel
          ÉG G G G G
   Z Z Z Z ÉZ Z Z Z     vowel
        H ÉH H         vowel
    T T T É
    V V V ÉV           vowel
          ÉA
      F F ÉF           vowel?
      X X É
        I ÉI           vowel?
          ÉU
```

The general principles are repeated. Vowels distinguish themselves from consonants as they are represented by

1) high frequency letters,
2) high frequency letters that do not contact each other
3) high frequency letters with great variety of contacts
4) high frequency letters with am affinity for low frequency PT consonants

Step 7: Prepare the partial enciphering alphabet and substitute into the cryptogram.

```
PT: 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
CT: S       Z       V         T H     D G F I
                                      F G
```

P-1 revisited and rewritten:

```
S F D Z F I O G H L P Z F G Z D Y S P F H B Z D S
a t r e t u   s o     e t s e r   a   t o   e r a
  s     s     t         s t           s

G V H T F U P L V D F G Y V J V F V H T G A D Z Z
s i o n t       i r t s   i   i t i o n s   r e e
t       s           s t         s       t

A I T Y D Z Y F Z J Z T G P T V T Z B D V F H T Z
  u n   r e   t e   e n s   n i n e   r i t o n e
              s         t                 s

D F X S B G I D Z Y V T X O I Y V T E F V M G Z Z
r t   a   s u r e   i n     u   i n   t     s e e
  s       t                           s     t

T H L L V X Z D F M H T Z A I T Y D Z Y B D V F H
n o     i   e r t   o n e   u n   r e     r i t o
              s

T Z D F K Z D Z Z J S X I S G Z Y G A V F S L G Z
n e r t   e r e e   a   u a s e   s   i t a   s e
      s                     t     t       s     t

D T H H T C D Z R S V T Y Z D O Z F F H T Z A I T
r n o o n   r e   a i n   e r   e t t o n e   u n
                                  s s

Y D Z Y G A V D G Z Z T K H I T Y Z Y S D Z G H U
  r e   s   i r s e e n   o u n   e   a r e s o
        t       t                           t

Z F Z T G U P G D I X W G H X A S R U Z D F U I D
e t e n s     s r u     s o     a     e r t   u r
  s     t     t         t                 s

E G H T V E A G X X
  s o n i     s
  t           t
```

I have left out the frequencies above the letters for editorial space only.

We can see from a first reading that PT words operations, nine prisoners, and afternoon come thru.
G = t, F = s, B = p, L =f.

Step 8: Complete the solution. Prepare the Ct/Pt Key Alphabets.

Message: As result of yesterdays operations by first division three hundred seventy nine prisoners captured including sixteen officers. One hundred prisoners were evacuated this afternoon, remainder less one hundred thirteen wounded are to be sent by truck to chambersburg tonight.

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

METHOD B: Probable Word Attack

"Patties" in the CM usually come with a tip and are generally shorter than the above example. The tip constitutes a probable PT word or phrase and we search the CT for a pattern of CT letters that exactly match the probable word. The choice of probable words is aided or limited by the number and positions of repeated letters. Repetitions may be patent (visible externally) or latent (made patent as a result of the analysis. For the example DIVISION with a repeated I or BATTALION with the reversible AT, TA help the cryptanalysis even though word divisions were removed. Friedman named what we call patterns as idiomorphs. [FRE2] gives many pattern lists for solution of substitution and Playfair ciphers. TEA computer program at the Crypto Drop Box is an automated pattern list to 20 words. [CAR1], [CAR2], [WAL1], [WAL2] give idiomorphic data. The process of superimposing the plain text word over the correct cipher text will effect the entry to the cryptogram. Other references include: [RAJ1], [RAJ2],[RAJ3], [RAJ4], [RAJ5], [HEMP], [LYNC].

SOLUTION OF ADDITIONAL CRYPTOGRAMS PRODUCED BY SAME COMPONENTS

Once the cryptogram has been solved and the keying alphabet reconstructed, subsequent messages which have been enciphered by the same means solve readily.

P- 2.

Suppose the following message is intercepted slightly later at the same station.

```
  I Y E W K   C E R N W   O F O S E   L F O O H   E A Z X X
```

P - 1. reconstruction and arbitrarily set at L = a.

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

     Cryptogram     I Y E W K    C E R N W
     Equivalents    P Y B F R    L B H E F
```

running down the sequence yields CLOSE YOURS as a generatrix.

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

Set the cipher component against the normal at C = i.

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

Solving: Close your station at two PM.

[FRE1] discusses keyword recovery processes on pp 85 -90. Also see [ACA].

VOWEL CIPHER

Louis Mansfield introduced the concept of a vowel substitution cipher in 1936. [MANS] In the vowel cipher the key alphabet is written into a square with j=i, like this:

```
                    COLUMN

                  A   E   I   O   U
                  É-----------------
              A Éa   b   c   d   e
                É
              E Éf   g   h   i/j k
                É
              I Él   m   n   o   p
   ROW          É
              O Éq   r   s   t   u
                É
              U Év   w   x   y   z
```

Enciphering is row by column or t = OO ; h = EI ; and e = AU.

The entire CT message enciphered is a succession of vowels. The CT will be exactly twice as long as the PT. This method is not more difficult to crack than the standard Aristocrat but our focus is on the frequency of vowel combinations in the CT. The PT equivalents do not have to be in standard order. Try this example.

VV-1.

```
  1      2              3                  4         5
OUOE    OUAE   OUIUIAAIUIUUOEOUAOAI    OEEOUUOE   OEEOAI

     6                 7            8
UOEUUEUEAIAEOE   OOAIOEUUOUUEAE   EEUO

  9          10             11              12       13
UUUEUE   OEUIEEEEIA   IUEEAOAIIUAIOAOEAE    IOAI   EIUUUI-

                          14           15       16
AIUOEUUEUEEA    EIEEIUIAOUUEAIOO   UUOAOO   AEAIOAOE

17       18            19
OEEE    EUAE     UIAIIIEUUEUUUIUEEA.
```

To solve this cipher we list the various combinations of two vowels.

```
    AA        EA  (2)    IA  (4)   OA  (3)    UA

    AE  (6)   EE  (6)    IE        OE  (11)   UE  (10)

    AI  (11)  EI  (2)    II  (1)   OI         UI  (5)

    AO  (2)   EO  (2)    IO  (1)   OO  (3)    UO  (3)

    AU        EU  (4)    IU  (4)   OU  (6)    UU  (7)
```

AI and OE appear 11 times and often as finals. Either might be the "e". OE appears as an initial. WE try OE as t and AI as e. Word 5 becomes 'the.' Word 4 confirms as 'that.' UU = a. UE = I. The normal procedure of test and confirm gives us the message: It is imperative that the fullest details of all troop movements be carefully compiled and sent to us regularly.

The partial keying square is:

	A	E	I	O	U
A		s	e	v	
E	y	o	c	h	u
I	p		g	b	m
O	n	t		d	i
U		l	r	f	a

MIRABEAU'S CIPHER

Comte de Mirabeau (1749 - 1791) was one of the great orators in the National Assembly, the body that governed France during the early phases of the French Revolution. He was a political enemy of Robespierre. He developed a simple substitution variant to relay his court messages to Louis XVI (who rejected his moderate advise). His father, the Marquis de Victor Riqueti Mirabeau imprisoned his son for failure to pay debts. He devised this system during his stay in debtors prison.

The Mirabeau system of ciphering letters of the alphabet are divided into five groups of five letters each. Each letter is numbered according to its position in the group. The group is also numbered. The key alphabet is arranged as follows:

```
 1 2 3 4 5      1 2 3 4 5      1 2 3 4 5
 I S U W B      K T D Q R      X L P A E
     6              8              4

           1 2 3 4 5    1 2 3 4 5
           G O Y V F    Z M C H N
               7            5
```

Encipherment of the phrase ' the boy ran' would be 82.54.45, 65.72.73, 85.44.55 where the t is referenced by group 8,letter2. Solution of messages is clearly by frequency analysis, the key being reconstructed from the message. Mirabeau experimented by reversing number order in this positional number system and adding nulls to confuse the interloper. One of the interesting complications added by Mirabeau was to express the CT as a fraction with group number as numerator and position number as denominator. Other figures were added to foil decipherment. In such a case the alphabet is grouped into fives as before, but the groups and positions are each numbered with the same five figures. So:

```
1 2 3 4 5      1 2 3 4 5      1 2 3 4 5
O A G P U      T C N H Y      X M F I S
    1              2              3

1 2 3 4 5      1 2 3 4 5
L Q W B V      R E K Z D
    4              5
```

Enciphering 'the boys run' :

```
2 2 5    4 1 2 3    5 1 2
- - -    - - - -    - - -
1 4 2    4 1 5 5    1 5 3
```

adding numerals 6, 7, 8, 9, 0 as non-values both above and below the line increase the security slightly. Or:

```
29  27  50     48  17  29  39     56  10  28
--  --  --     --  --  --  --     --  --  --
71  64  92     74  94  85  65     91  75  83
```

The recipient reads the message by cancelling the non-values and using the others. A key to recognition of this cipher is that the non-values (nulls) are never employed as group or position numbers.

The complicated form of the Mirabeau is solved by preparing a Fractional Bigram sheet and reducing out the non-values. Suppose we encipher the phrase 'we have been here':

```
2  1  4  5      4  5  5  2      1  5  5  5
-  -  -  -      -  -  -  -      -  -  -  -
4  2  5  2      4  2  2  3      4  2  1  2
```

Using non-values (6,7,8,9,0) as:

```
62  10  40  65    47  57  75  62    27  58  57  85
--  --  --  --    --  --  --  --    --  --  --  --
48  62  95  20    84  27  92  30    49  62  19  29
```

The five 'e' s which occur are different each time.

```
65  57  75  58  85
--  --  --  --  --
20  27  92  62  29
```

The fractional group sheet proceeds like a Bigram analysis. Instead of letters we use fractions.

The first fraction would be noted in four different ways, e.g.,

```
6  6  2  2
-  -  -  -
4  8  4  8
```

```
          65                      6  6  5  5
the group --    would be catalogued -  -  -  -
          20                      2  0  2  0
```

```
           5
The fraction -  (which is the real e) will eventually assume
           2
```

its normal frequency and thus display its identity. Armed with the fact that 5/2 represents e, we cancel out all the non-values which occur with this fraction. Each time we cancel out a non-value, we do so for the entire cryptogram. Even if the 5/2 represents another letter, such as t, the uniliteral frequency distribution will be present in the CT.

TELEPHONE CIPHER VARIATION - CHARLES SCHWAB

Hardly a cipher, but a modern substitution system effecting 10 million brokerage customers is the Charles Schwab Telephone Automated Customer Service System. The telephone is used for the enciphering of literal and numerical data to the Schwab computer system. So:

```
1  -
2  -  A B C
3  -  D E F
4  -  G H I
5  -  J K L
6  -  M N O
7  -  P R S          NO Q     use 99
8  -  T U V
9  -  W X Y          NO Z     use 98
*  -
0  -
#  -
```

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

STRIKE PRICE CODES

```
 A   B    C    D    E    F    G    H    I    J    K    L

 5   10   15   20   25   30   35   40   45   50   55   60
105  110  115 120  125  130   135  140  145 150  155  160
205  210  215 220  225  230   235  240  245 250  255  260
305  310  315 320  325  330   335  340  345 350  355  360

 M   N    O    P    Q    R    S    T    U    V    W    X

65   70   75   80   85   90   95   100   7.5 12.5 17.5 22.5
165  170  175 180  185  190   195  200
265  270  275 280  285  290   295  300
365  370  375 380  385  390   395  400
```

Other combinations of the above indicate special actions. 10 = accept. 90 = reject. * = return, end, # - account terminator. Note that 1 number represents 3 equivalents. Schwab uses the position to indicate the letter similar to the ancient Masonic Cipher. For example, Janus Enterprise Fund symbol is JAENX = 51-21-32-62-92.

An order to buy 750 shares JAENX at a limit price of 23.5 plus a MAY call for 100 shares of BAX at strike price 25 might include the following entries in the electronic order:

18002724922, 61702554#, xxxx#, 1, 1, 750, 5121326292, 54,

23*50, *, 1, 1, 100, 222192, 32, 32, 10, *

which represents the telephone number of Charles Schwab, account number and PIN, order codes, limit code, transfer codes, menu response items. Other codes would allow you to move around your account and monitor the order. [Schw]

Note also that the basis telephone code is a 12 by 3 matrix.

```
        1 2 3
    1   - b b b
    2   - A B C
    3   - D E F
    4   - G H I
    5   - J K L     b - represents available
    6   - M N O         information slot
    7   - P R S
    8   - T U V     i.e. 9 = W X Y, but 93 = Y
    9   - W X Y                           only
10  *   - b b b
11  0   - b b b
12  #   - b b b
```

It is easy to see how this process could be expanded to larger and larger keyspaces. See references [BOSW], [KOBL] and [WEL]. for a fair discussions of the numerical requirements involved. A good discussion of the Information Theory is found in reference [RHEE]. A look at modern design criteria for bank fund transfer and similar PIN systems in found in Meyer and Matyas. [MM]

MORE COMPUTER AIDS

Dr. Caxton C. Foster who wrote "Cryptanalysis for Micro-computers, while at the University of Massachusetts, has generously donated his computer programs on substitution and transposition to the class. I have sent an updated disk to our CDB. [CCF] GWEGG has Cryptodyct on disk written in DbaseIV. Contact him for a copy.

A review of the entire field of applied cryptography is presented in Bruce Schneier's book. Most of the material is beyond the scope of this class, however a PC source / program diskette is included with his book. There are ITAR limitations associated with his disk. We will cover some of the historic symmetric algorithms such as Vigenere and Playfair ciphers. [SCHE]

HOMEWORK ASSIGNMENTS

Pd-1. Daniel

H Z K L X A L H X P N C I N Z X F L I X G N W Q X P N Z K T L N K X O

L X N I Z X G I N X P N E Z K X W Q X P Z X L H X P N C I N Z X S N Q

N T X W Q X P N W V S N I K L K H B L X N W Q L X H F Z I L N X A Z K

S B W E N I.

Pd-2. Join the army. Daniel

F L B B A O I A F Q E A O M Z U I L O N R Z O Q A O P I L O M O L S F

P F L I P F L B B A O E R I C A O Q E F O P Q B L O W A V H Z O W E A

P X Z Q Q G A P Z I V V A Z Q E G A Q E F H T E L G L S A P L R O W L

R I Q O U F I E F P E A Z O Q Z I V I L Q T F Q E E F P G F M P L I G

U B L G G L T H A.

SOLUTION TO HOMEWORK PROBLEMS FROM LECTURE 1

First assume that only English is involved in all 5 problems. (This may not be true third round.) My thanks to both WALRUS and SNAIL PACE for detailed solutions.

Problem A-1. can be solved by the Pattern Word method.

A-1. Bad design. K2 (91) AURION

```
   1        2          3          4      5       6
V G S   E U L Z K   W U F G Z   G O N   G M   V D G X Z A J U =

                  7           8      9       10
X U V B Z    H B U K N D W   V O N   D K   X D K U H H G D F =

        11     12       13          14
  U K   Y D K   V G U N   A J U X O U B B S

       15          16     17        18
X D K K G B P Z K   D F   N Y Z    B U L Z .
```

Lots of two and three letter words. One of the three letter words most likely 'the'. The GON GM combination is a possible wedge. Note that N = 6/91 = 6.6% of CT, Z = 8/91 = 8.9% and V = 5/91 = 5.5% of CT. Word 10 pattern is (556) 291 on 12L word = disappointed. Note the 'ted' ending fits with word 17 t_e = the. so Y=h. Word 16 confirms DF = in. Plugging in A-1, we have:

A-1. Bad design. K2 (91) AURION

```
   1        2          3          4      5       6
b o y    a   e s    a n o e   o u t   o    b i o d e g r a
V G S   E U L Z K   W U F G Z   G O N   G M   V D G X Z A J U =
                  7           8      9       10
d a b l e    p l a   t i    b u t   i s   d i s a p p o i n
X U V B Z    H B U K N D W   V O N   D K   X D K U H H G D F =
        11     12       13          14
t e d   a s   h i s   b o a t   g r a d u a l l y
N Z X   U K   Y D K   V G U N   A J U X O U B B S
       15          16     17        18
d i s s o l   e s   i n   t h e    l a   e
X D K K G B P Z K   D F   N Y Z    B U L Z .
```

Words 11, 12 show: as his. Word 6 looks like biodegradable, and V confirms as a b in boat (word 13). Word 8 becomes but. Word 14 ends ally. The messages reads: Boy makes canoe out of biodegradable plastic but is disappointed as his boat gradually dissolves in the lake. The keyword recovered is MAYDAYCALL. Note that the tip was useless. One 12 letter pattern word opens her up like a clam.

A-2. Not now. K1 (92) BRASSPOUNDER

```
K D C Y   L Q Z K T L J Q X   C Y   M D B C Y J Q L :   " T R   H Y D    F K X C ,

F Q   M K X   R L Q Q I Q   H Y D L   M K L   D X C T W   R D C D L Q

J Q M N K X T M B   P T B M Y E Q L   K   F K H   C Y   L Q Z K T L   T C . "
```

Solve A-2 by eyball method. One letter word = a, look for you .. your combination, the word to in the first four words before quotation. The message reads: Auto repairmen to customer: " If you want, we can freeze your car until future mechanics discover a way to repair it."

A-3. Ms. Packman really works! K4 (101) APEX DX

```
* Z D D Y Y D Q T   Q M A R P A C ,   * Q A K C M K   * T D V S V K .   B P   W V G

Q N V O M C M V B :   L D X V   K Q A M S P D   L V Q U ,  L D B Z I   U V K Q F

P O   W A M U X V ,   E M U V P   X Q N V ,  U A M O Z   N Q K L M O V

( S A P Z V O ) .
```

APEX DX always sends an interesting con. Vowel splitting yields V, D, M, Q, P. Word 7 suggests that M=i. Words 15 and 16 look like video game, the word fridge comes to bear. The message reads: Kuujjuaq airport, Arctic Quebec. So few amenities: huge caribou head, husky decal on fridge, video game, drink machine (broken). Kw = chimo; FORT.

A-4. Money value. K4 (80) PETROUSHKA

```
D V T U W E F S Y Z   C V S H W B D X P   U Y T C Q P V   E V Z F D A   E S T U W X

Q V S P F D B Y   P Q Y V D A F S ,   H Y B P Q   P F Y V C D   Q S F I T X

P X B J D H W Y Z .
```

Using the consonant-line method:

```
 CEHZAUIXP
 -----------
 TTTT ÉT        VOWEL
  VV  ÉVVV      VOWEL
  YY  ÉYY
  Q   ÉQQQ
DDDD  ÉDD
 WW   ÉWWWW
  SS  ÉS
 F    ÉFFFFF     VOWEL
    BBÉB
    P ÉP
      É
```

J can be wrongly assumed to be a consonant. Digraph HW and rt /tr reversal fails but st/ts reversal gives information. The word merchants can be found in a non-pattern word list. The ch combination fits CT HW. The message reads: Neighborly merchants glimpse beyond bright personal splendor, clasp solemn profit staunchly. Kw(s)= sprightly; BEHAVIOR.

A-5. Zoology lesson. K4 (78) MICROPOD

```
A S P D G U L W ,   J Y C R   S K U Q   N B H Y Q I   X S P I N

O C B Z A Y W N = O G S J Q   O S R Y U W ,   J N Y X U   O B Z A   ( B C W S

D U R B C )   T B G A W   U Q E S L .   * C B S W
```

Note the entry B C W S... *C B S W. Try also.. LAOS. The consonant line yields S, U, Y, B as vowels. The message is; Koupreys, wild oxen having tough blackish=brown bodies, white back (also pedal) marks enjoy Laos. Kws = undomestic; BOVINE.

REFERENCES / RESOURCES

[ACA] ACA and You, Handbook For Members of the American Cryptogram Association, 1995.

[BARK] Barker, Wayne G., "Cryptanalysis of The Simple Substitution Cipher with Word Divisions," Aegean Park Press, Laguna Hills, CA. 1973.

[BAR1] Barker, Wayne G., "Course No 201, Cryptanalysis of The Simple Substitution Cipher with Word Divisions," Aegean Park Press, Laguna Hills, CA. 1975.

[BOSW] Bosworth, Bruce, "Codes, Ciphers and Computers: An Introduction to Information Security," Hayden Books, Rochelle Park, NJ, 1990.

[B201] Barker, Wayne G., "Cryptanalysis of The Simple Substitution Cipher with Word Divisions," Course #201, Aegean Park Press, Laguna Hills, CA. 1982.

[BP82] Beker, H., and Piper, F., " Cipher Systems, The Protection of Communications", John Wiley and Sons, NY, 1982.

[CAR1] Carlisle, Sheila. Pattern Words: Three to Eight Letters n Length, Aegean Park Press, Laguna Hills, CA 92654, 1986.

[CAR2] Carlisle, Sheila. Pattern Words: Nine Letters in Length, Aegean Park Press, Laguna Hills, CA 92654, 1986.

[CCF] Foster, C. C., "Cryptanalysis for Microcomputers", Hayden Books, Rochelle Park, NJ, 1990.

[DAGA] D'agapeyeff, Alexander, "Codes and Ciphers," Oxford University Press, London, 1974.

[DAN] Daniel, Robert E., "Elementary Cryptanalysis: Cryptography For Fun," Cryptiquotes, Seattle, WA., 1979.

[DOW] Dow, Don. L., "Crypto-Mania, Version 3.0", Box 1111, Nashua, NH. 03061-1111, (603) 880-6472, Cost $15 for registered version and available as shareware under CRYPTM.zip on CIS or zipnet.

[ELCY] Gaines, Helen Fouche, Cryptanalysis, Dover, New York, 1956.

[EPST] Epstein, Sam and Beryl, "The First Book of Codes and Ciphers," Ambassador Books, Toronto, Canada, 1956.

[GIVI] Givierge, General Marcel, " Course In Cryptography," Aegean Park Press, Laguna Hills, CA, 1978.

[GODD] Goddard, Eldridge and Thelma, "Cryptodyct," Marion, Iowa, 1976

[GORD] Gordon, Cyrus H., " Forgotten Scripts: Their Ongoing Discovery and Decipherment," Basic Books, New York, 1982.

[FR1] Friedman, William F. and Callimahos, Lambros D., Military Cryptanalytics Part I - Volume 1, Aegean Park Press, Laguna Hills, CA, 1985.

[FR2] Friedman, William F. and Callimahos, Lambros D., Military Cryptanalytics Part I - Volume 2, Aegean Park Press, Laguna Hills, CA, 1985.

[FR3] Friedman, William F. and Callimahos, Lambros D., Military Cryptanalytics Part III, Aegean Park Press, Laguna Hills, CA, 1995.

[FR4] Friedman, William F. and Callimahos, Lambros D., Military Cryptanalytics Part IV, Aegean Park Press, Laguna Hills, CA, 1995.

[FRE] Friedman, William F. , "Elements of Cryptanalysis," Aegean Park Press, Laguna Hills, CA, 1976.

[FR22] Friedman, William F., The Index of Coincidence and Its Applications In Cryptography, Publication 22, The Riverbank Publications, Aegean Park Press, Laguna Hills, CA, 1979.

[HA] Hahn, Karl, " Frequency of Letters", English Letter Usage Statistics using as a sample, "A Tale of Two Cities" by Charles Dickens, Usenet SCI.Crypt, 4 Aug 1994.

[HEMP] Hempfner, Philip and Tania, "Pattern Word List For Divided and Undivided Cryptograms," unpublished manuscript, 1984.

[INDE] PHOENIX, Index to the Cryptogram: 1932-1993, ACA, 1994.

[KOBL] Koblitz, Neal, " A Course in Number Theory and Cryptography, 2nd Ed, Springer-Verlag, New York, 1994.

[KULL] Kullback, Solomon, Statistical Methods in Cryptanalysis, Agean Park Press, Laguna Hills, Ca. 1976

[LAFF] Laffin, John, "Codes and Ciphers: Secret Writing Through The Ages," Abelard-Schuman, London, 1973.

[LYNC] Lynch, Frederick D., "Pattern Word List, Vol 1.," Aegean Park Press, Laguna Hills, CA, 1977.

[MANS] Mansfield, Louis C. S., "The Solution of Codes and Ciphers", Alexander Maclehose & Co., London, 1936.

[MM] Meyer, C. H., and Matyas, S. M., " CRYPTOGRAPHY - A New Dimension in Computer Data Security, " Wiley Interscience, New York, 1982.

[NIC1] Nichols, Randall K., "Xeno Data on 10 Different Languages," ACA-L, August 18, 1995.

[NIC2] Nichols, Randall K., "Chinese Cryptography Part 1," ACA-L, August 24, 1995.

[OP20] "Course in Cryptanalysis," OP-20-G', Navy Department, Office of Chief of Naval Operations, Washington, 1941.

[PIER] Pierce, Clayton C., "Cryptoprivacy", 325 Carol Drive, Ventura, Ca. 93003.

[RAJ1] "Pattern and Non Pattern Words of 2 to 6 Letters," G & C. Merriam Co., Norman, OK. 1977.

[RAJ2] "Pattern and Non Pattern Words of 7 to 8 Letters," G & C. Merriam Co., Norman, OK. 1980.

[RAJ3] "Pattern and Non Pattern Words of 9 to 10 Letters," G & C. Merriam Co., Norman, OK. 1981.

[RAJ4] "Non Pattern Words of 3 to 14 Letters," RAJA Books, Norman, OK. 1982.

[RAJ5] "Pattern and Non Pattern Words of 10 Letters," G & C. Merriam Co., Norman, OK. 1982.

[RHEE] Rhee, Man Young, "Cryptography and Secure Communications," McGraw Hill Co, 1994

[ROBO] NYPHO, The Cryptogram, Dec 1940, Feb, 1941.

[SCHN] Schneier, Bruce, "Applied Cryptography: Protocols, Algorithms, and Source Code C," John Wiley and Sons, 1994.

[SINK] Sinkov, Abraham, "Elementary Cryptanalysis", The Mathematical Association of America, NYU, 1966.

[SMIT] Smith, Laurence D., "Cryptography, the Science of Secret Writing," Dover, NY, 1943.

[STIN] Stinson, D. R., "Cryptography, Theory and Practice," CRC Press, London, 1995.

[SCHW] Schwab, Charles, "The Equalizer," Charles Schwab, San Francisco, 1994.

[TUCK] Harris, Frances A., "Solving Simple Substitution Ciphers," ACA, 1959.

[WAL1] Wallace, Robert W. Pattern Words: Ten Letters and Eleven Letters in Length, Aegean Park Press, Laguna Hills, CA 92654, 1993.

[WAL2] Wallace, Robert W. Pattern Words: Twelve Letters and Greater in Length, Aegean Park Press, Laguna Hills, CA 92654, 1993.

[WEL] Welsh, Dominic, "Codes and Cryptography," Oxford Science Publications, New York, 1993.

[WRIX] Wrixon, Fred B. "Codes, Ciphers and Secret Languages," Crown Publishers, New York, 1990.

[ZIM] Zim, Herbert S., "Codes and Secret Writing." William Morrow Co., New York, 1948.

LECTURE 3 OUTLINE

I expect to cover the following subjects in my next lecture:

Variant Substitution Systems

- Simple Numerical Ciphers
- Multiliteral Substitution with Single Equivalent Cipher Alphabets
- Baconian Cipher
- Hayes Cipher
- Trithemian Cipher
- Other historical variants.

LECTURE 4

We will cover recognition and solution of XENOCRYPTS (language substitution ciphers) in detail.

LECTURE 1 ERRATA

The Parker Hitt distribution of letters is per 20,000 letters. The phrase "aa a" should have been "as a". I will correct others that I have been advised of and retransmit them to our CDB.

CLASS NOTES

Our class seems to have leveled off at 86 students! This may be a record size for any public cryptography class offered to date. I thank you for your confidence.

NORTH DECODER, in addition to running the ACA-L list server and Crypto Drop Box superbly, has taken it upon himself to act as my grammarian. I appreciate his help finding the late night "additions/subtractions." TATTERS has volunteered as an assistant with LEDGE. Thank you.

TATTERS, in addition to making available his microcomputer crypto programs to the class has agreed to assist on the Cipher Exchange lectures at the beginning of 1996. Thank you. LEDGE will be assisting on the Cryptarithms Lectures. Thank you. My typing fingers thank you both!