SRH- 214

ELEMENTARY COURSE IN CRYPTANALYSIS

OFFICE OF CHIEF OF NAVAL OPERATIONS
NAVY DEPARTMENT

CIRCA 1939

REVIEWER'S NOTE:

The first review of this document was conducted by personnel of the U. S. Navy. The original classified versions were retained by them and have been placed in the NSG Repository, Crane, Indiana

N.B.:

This document is very similar to that which has been issued as SRH-216. There were enough differences , however, to warrant issuance under a separate number.

CONTENTS

~~RESTRICTED~~

NAVY DEPARTMENT
Office of Chief of Naval Operations
WASHINGTON

0993

*

ELEMENTARY COURSE IN CRYPTANALYSIS

- ASSIGNMENT No.1 -

- INTRODUCTION -

1. The student should immediately grasp the idea that his success as a cryptanalyst will depend almost entirely upon his own initiative and industry. A complete mastery of the art of cryptanalysis can only come as the result of independent study and the solution of cryptograms themselves.

2. Although the principles to be dealt with in elementary cryptanalysis are widely known, a knowledge of the mere existence of this course must be restricted to members of the Naval Service. The number of persons skilled in cryptanalytics, their identity, and their degree of proficiency, must be carefully guarded. Do not discuss anything connected with this course with anyone outside the Navy or Naval Reserve.

- HISTORICAL NOTES -

3. Cryptography, in its simpler forms, would appear from the evidence available to be as old as the written language itself. In fact, it seems probable that it may have in some instances actually ante-dated the written language, for we find numerous indications of usage, in the most remote times, of arbitrary signs for conveying secret information. Certainly by the time of the Greek and Roman civilizations we find cryptography occupying an important place in practically all important military operations. Julius Caesar is reported to have used a system in which each letter was replaced by the letter in the alphabet in the third position from it, such as, D for A, etc., while Augustus used the letter preceding the desired letter. It is interesting to note that this system with variations is still in use by amateur correspondents today.

4. Throughout the Middle Ages the art and practice of cryptography continued to develop. Numerous scholars and philosophers attempted to construct a perfect cipher. Among these might be mentioned Francis Bacon and Blaise de Vigenere, both of whom contributed materially to the art without, however, achieving the ideal for which they sought.

5. In these early days the transmission of communications was ordinarily restricted to the use of couriers and other equally slow and uncertain means. Frequently, the use of trustworthy messengers achieved the result desired and the employment of cryptography was not always essential to secrecy. With the advent of telegraphy all this was changed. Communication became almost instantaneous, but the channels themselves could not be so thoroughly guarded. Wire tapping was nearly always possible with the ordinary telegraph or cable, but with the advent of radio, ev this became unnecessary, for radio transmissions are always ava able to anyone with a sufficiently sensitive receiver. All t s tended strongly to concentrate attention on cryptography as the only means available whereby a reasonable degree of secrecy could be attained, and led to a much more rapid advance in the art. It also served to crystalize development along those lines which were suitable to telegraphic transmission, eliminating to a large extent the importance of secret inks, as well as pictorial, and such other kindred methods with which we need not concern outselves. A cryptogram to be transmitted by telegraphic means must, of necessity, consist primarily of letters or numerals whether alone or in combination.

6. While the history of the development of cryptography is none too complete, the history of cryptanalysis is even more fragmentary and one must resort even more to surmise. It is likely that cryptanalysis is as old as cryptography itself, for it seems to be an innate trait of human nature to attempt to read the secrets of others. Fortunately for the peace of mind of the majority of us, this trait seems to have been most often deflected into the pursuit of the puzzles and riddles which have occupied mankind in all ages.

7. Despite the general lack of historical material numerous instances of the use of cryptanalysis do stand out. After the battle of Naseby, Cromwell employed the English mathematician, John Wallis, to decipher the secret papers of Charles I, proving conclusively that the King had been guilty of double dealing in his negotiations. Another early investigator, Francois Viete, successfully analyzed the cipher used by the Holy League, but the effort very nearly cost his life, for it was charged that only by the use of necromancy could he have obtained the key, and it was with great difficulty that he cleared himself. At a much later

date, Edgar Allen Poe delighted the world with "The Gold Bug" and his treatises on cipher analysis. Also, much reference to cryptanalysis is to be found in modern detective literature, but, in general, history is strangely silent on this important subject. This is no doubt largely due to the high degree of secrecy with which such matters have of necessity always been clothed. Disclosures of any kind are highly inimical to the interests of the military or diplomatic cryptanalyst, as well as to the country which he serves, for such disclosures almost invariably close an important avenue of information. It should not be concluded from this however, that cryptanalysis has failed to play an important part, both in peace and war. An instance of this may be noted in the affair of the Zimmermann note to Mexico during the World War. The details of that affair are so well known that they need not be rehearsed here, but we should note how the reading of a single enemy message so materially aided England in bringing the United States into the war. In the more restricted fields of military strategy and tactics, it is quite obvious that the commander who has full knowledge of the enemy's plans and intentions through the reading of his intercepted despatches is in a much better position for bringing the action to a successful conclusion than one who is denied this information. Thus the military importance of the successful cryptanalyst can scarcely be overemphasized.

8. The rise of modern communication methods, especially radio, have had two very profound effects on cryptanalysis. Due to the resultant improvement of cryptographic methods noted above, the skill and labor involved in the processes of analytical solution has been greatly increased. On the other hand, however, there has been placed in the hands of the cryptanalyst an almost infallible source of cryptographic material which in former times could scarcely be obtained except as the result of fortuitous chance. This has led to the development of cryptanalysis to the high status which it holds almost universally today. With this development, regrettably enough, the United States has scarcely kept pace. It is doubtful if the time will ever come when this country can and will maintain in times of peace a highly developed and well organized cipher bureau such as are reputedly maintained by other countries and for that reason the primary reliance in time of war must be placed on the skilled amateur cryptanalyst. It is in the hope of establishing such a body of trained amateurs that this course has been inaugurated.

DEFINITIONS

9. The definitions found in this course have been taken from the Army Extension Course in "Elementary Military Cryptography" through the courtesy of Major W. F. Friedman, Signal Reserve, U.S. Army.

10. Cryptology is that branch of knowledge which treats of all the means and methods of secret intercommunication.

11. Cryptography is that branch of cryptology which treats of the various means, methods, and devices for converting plain-text messages into cryptograms and reconverting the so-produced cryptograms into their plain-text form by a direct reversal of the steps or processes employed in the original conversion.

12. Plain text is writing which conveys an intelligible meaning in the language in which it is written.

13. Cryptographic text is writing which conveys no intelligible meaning in any language, or which apparently conveys an intelligible meaning that is not the real meaning intended to be conveyed.

14. A cryptogram is a communication written in secret language, which may be transmitted by any of the agencies of intercommunication. As mentioned before, we are concerned only with cryptograms which can be transmitted by radio or telegraph.

15. Cryptographing and decryptographing are accomplished by means collectively designated as codes and ciphers. In cipher systems cryptograms are produced by applying the cryptographic treatment to individual letters of the plain text messages, whereas in code systems cryptograms are produced by applying the cryptographic treatment to entire words, phrases, and sentences of the plain-text messages. The code systems become, in the final analysis, a more or less highly specialized form of substitution.

16. Substitution and transposition are the only two distinctly different types of treatment which may be applied to plain text to convert it into secret text, yielding two different classes of cryptograms. In substitution the elements of the plain text retain their original positions or sequences, but are replaced by other elements with different values or meanings. In transposition the elements or units of the plain text, whether one is dealing with individual letters or groups of letters, retain their original identities but merely undergo some change in their relative positions or sequences so that the message becomes unintelligible.

17. It may be stated that, as a general rule, all or nearly all cryptographic systems suitable for practical use can be broken down, or solved, that is, properly prepared cryptograms can be "translated" or read without a knowledge or possession of a general cryptographic system and the specific key applying to the cryptograms.

18. That branch of cryptology which deals with the principles, methods, and means employed in the solution or analysis of cryptograms is called cryptanalytics.

19. The steps and operations performed in applying the principles of cryptanalytics constitute cryptanalysis. To cryptanalyze or to decrypt a cryptogram is to solve it by cryptanalysis.

20. The normal alphabet for any language is one in which the sequences of sounds or symbals have been definitely fixed by long usage or convention.

21. A cipher alphabet is one in which the elementary speech-sounds are represented by characters other than those represenging them in the normal alphabet.

22. When the plain text of a message is converted into secret text by the use of one or more cipher alphabets, the resultant cryptogram constitutes a substitution cipher.

23. It will be convenient to designate that component of a cipher alphabet constituting the sequences of speech-sounds, the plain component, and the component constituting the sequence of symbols, the cipher component.

24. As regards the sequence of the letters forming its cipher component, cipher alphabets are of two kinds:

(a) Standard cipher alphabets, in which the sequence of letters in the cipher component is the same as the normal, but either shifted from its normal point of coincidence with the plain component or reversed in direction.

Examples -

Direct Standard Cipher Alphabet

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

It is obvious that the cipher component can be applied to the plain component at any one of 26 points of coincidence (except the one which coincides exactly).

Reversed Standard Cipher Alphabet

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

Here the cipher component can be applied to the plain component at any one of 26 points of coincidence. This is also an example of a reciprocal alphabet, that is, the equivalents are reversible or reciprocal in pairs. (A plain is Q cipher, and Q plain is A cipher). Thus reciprocal alphabets may serve either as enciphering or deciphering alphabets.

(a) Mixed cipher alphabets, in which the sequence of letters or characters in the cipher component is no longer the same as the normal in its entirety.

Example -

Random Mixed Cipher Alphabet

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

Systematically Mixed Cipher Alphabet

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

Systematically-mixed cipher alphabets will be discussed in Assignment No.2.

25. If a cipher alphabet is drawn up and a message is enciphered by its means, letter-for-letter consistently throughout the message, it is said that the cryptogram has been enciphered by a single alphabet, and it is a single-alphabet substitution cipher. When only one alphabet is employed, the system is technically called monoalphabetic substitution, and when two or more cipher alphabets are employed, it is called polyalphabetic substitution.

26. The following problem is an example of a mono-alphabet cipher of the simplest type, that is, one of which the plain language word lengths have been left intact, and not combined in 5-letter groups for telegraphic transmission, as is ordinarily done:

EIFIXQZS QXXQOJM PMZDM DCXILIMN PS IXQFIQD DQVQF
QWXJCZIXEMY XJQX KMZEQD YWPEQZIDMY JQN PMMD CZNMZMN
XC KC IDXC EMNIXMZZQDMQD YMQ XC QXXQOG SJM XUC
XMQDYBCZXY OQZZSIDK QEMZIOQD YCFNIMZY EQZYMIFFM QDN
XCWFCD EMZICN ZMAWMYX FCDNCD PM IDLCZEMN HQOGYCD

27. The basic principle of cipher solution is that underlying the cipher text is plain text and the peculiarities of the plain text language itself lead to the solution. Solution is thus based on language itself rather than on the frequency of occurrence of individual letters. To fix this principle firmly in the mind of the beginner, an illustration of solution of this problem is given.

28. First, the cipher text is examined for repetitions and peculiarities of letter distribution. Repetitions have been underlined, and they represent words or parts of words which are probably common in English, otherwise they would not be repeated in such a short message. Peculiar letter distributions are: doubled letters, repeated letters within a small number of letters, and reversed digraphs. Some of the peculiar distributions have been overlined in the cryptogram.

29. It should be remembered here that a language cannot be written or spoken without using certain connective words, syllables, and phrases. The most common of these are: that, which, tion, ing, ence, the, been, have, had, has, and, to, of, but, not, in. Also, punctuation is often used so the words "period", "comma", and "stop" may be added to the list. Since these words appear so often in the English language, regardless of subject matter, one or more of them has an excellent chance of appearing as a repetition in the cipher text.

30. Having carefully scrutinized the text, the next step is to make assumption of plain language values. André Langie, a French author of works on Cryptanalysis, has said that the motto of the cryptanalyst should be: "Let's suppose". He has also said

that the most important aid in cipher solution is a good eraser. In other words, make logical assumptions where possible; if they do not lead to solution, erase the assumptions which have been proved incorrect and make others.

31. In our problem the word XJQX immediately attracts attention. First, we know it to be a complete word. Even when the text is not spaced into proper word lengths, such a combination would invite attention because it fits the very common word "that". Therefore, tentatively substitute T,H,A,T (plain) for XJQX. To verify this assumption, substitute the assumed values throughout the cipher text wherever X,J,Q, appear. (The student should follow through on this solution by actually performing each step). The assumption is certainly now a good possibility because of the excellent combinations which it gives elsewhere in the message: XJM (cipher) = TH - (plain); *JQN(c) = HA - (p); QXXQOG(c) = ATTA -- (p). If the initial assumption was correct, then the M of XJM must represent E(p) to make XJM(c) = THE. Also, in JQN(c) = HA -(p), N(c), probably represents either S(p) or D(p) to make HAS or HAD. Where XC(c) = T - (p), C(c) must represent O(p). Therefore, throughout the text substitute E(p) for M(c) and O(p) for C(c). This substitution gives some excellent combinations of plain letters and no combinations which are impossible. Look at IDXC(c) = -- TO(p). Obviously ID(c) = IN(p). Again substitute throughout. Now, there can no longer be any doubt as to the correctness of out initial assumption. IXQFIQD(c) = ITA - IAN, so F(c) = L(p); KC(c) = - O(p), So K(c) = G(p); QDN(c) = AN - (p), So N(c) = D(p). Substitute the newly recovered values, and continue the process. The entire cipher message is solved very easily from this point on.

32. The complete translation is: MILITARY ATTACHE BERNE NOTIFIED BY ITALIAN NAVAL AUTHORITIES THAT GERMAN SUBMARINES HAD BEEN ORDERED TO GO INTO MEDITERRANEAN SEA TO ATTACK THE TWO TRANSPORTS CARRYING AMERICAN SOLDIERS MARSEILLE AND TOULON PERIOD RE QUEST LONDON BE INFORMED JACKSON.

33. The problem was solved without paying any attention whatever to frequency tables, and without any knowledge whatever as to the nature of the text, except that it was in English. Only one assumption had to be made and then step by step it was only necessary to substitute obvious values after substituting the initial assumed values. Had the initial assumption been incorrect, it would have been erased and a new start made. There were other obvious breaks which would inevitably and soon be found by

* XJM (cipher) TH -(plain) will hereafter appear XJM(c) = TH - (P). (K) is also used to mean (key).

"trial and error" or, if you prefer, by hypothesis and test. The three words in sequence XC KC IDXC is an excellent starting point and would soon have been assumed to be TO GO INTO. Another was the two words QXXQOJM and QXXQOG. The latter would sooner or later have been assumed to be ATTACK and this would make the first ATTACHE.

34. The problems given the student in Assignment No. 1 for solution are to be solved in the same manner, which is called "by inspection". No frequency tables are to be employed. After solution of these problems, the student will readily see that ciphers of this type prove to be a very inadequate form of camouflage.

35. The lesson to be learned from Assignment No.1, which should never be forgotten in Cryptanalysis, is "The fundamental principle of cipher solution is based upon the peculiarities of the plain language itself".

ELEMENTARY
COURSE IN CRYPTANALYSIS

- ASSIGNMENT No.1 -

PART II

Answer the following questions:

1. What is the difference between cryptography and cryptanalysis? In cryptography, cryptograms are decrypted only by the direct reverse of the encrypting methods, whereas any method or means of decrypting is used in cryptanalysis.

2. What is the difference between a code and a cipher? The basic unit for enciphering and deciphering is the letter. The basic unit for encoding or decoding is a word, phrase or sentence.

3. How are substitution ciphers distinguished from transposition ciphers? Letters of the decrypted message will have the same frequency as a normal message in a transposition cipher but not in a substitution cipher.

- Solve The Following Problems -

PROBLEM No.1

Non-Naval Text

F T U E E T A D F O D K B F A S D M Y U E S U H Q Z

THIS SHORT CRYPTOGRAM IS GIVEN

M E M Z Q J Q D O U E Q U Z F T Q E A X G F U A Z

AS AN EXERCISE IN THE SOLUTION

A R M O U B T Q D N K U Z E B Q O F U A Z

OF A CIPHER BY INSPECTION

PROBLEM No.2

Non-Naval Text

D S V M Z H P V W Z Y L F G S R H K O Z M L U

WHEN ASKED ABOUT HIS PLAN OF

X Z H K Z R T M T V M V I Z O H G L M V D Z O O

CAMPAIGN GENERAL STONEWALL

Q Z X P H L M I V K O R V W G L Z M R M J F R H R G R E V

JACKSON REPLIED TO AN INQUISITIVE

X S Z K O Z R M X Z M B L F P V V K Z H V X I V G

CHAPLAIN CAN YOU KEEP A SECRET

B V H G S V V Z T V I X O V I R X Z M H D V I V W

YES THE EAGER CLERIC ANSWERED

D V O O H L X Z M R H Z R W G S V T V M V I Z O

WELL SO CAN I SAID THE GENERAL

PROBLEM No.3

Non-Naval Text

D C L C V S R U C T N S B U C G T O B S P P G R Y D

DEVELOPMENT OF MEANS FOR RAPID

E S U U M T Y E G N Y S T Z G O D S T C E S T O Y D C P G F V C

COMMUNICATION HAS DONE CONSIDERABLE

N S Z G P D E P C G N Y T A G F C N N C P

TOWARD CREATING A BETTER

Y T N C P T G N Y S T G V M T D C P O N G T D Y T A

INTERNATIONAL UNDERSTANDING

PROBLEM No.4

Non-Naval Text

XA ERK VZZB ZFB BQWO LZIBO
IN THE GOOD OLD DAYS WORDS
LKIK MQIKCUFFW OHQMKB XA
WERE CAREFULLY SPACED IN
MIWHEZVIQYO FXJK ERXO NUE
CRYPTOGRAMS LIKE THIS BUT
EZBQW LZIBO QIK IUA EZVKERKI
TODAY WORDS ARE RUN TOGETHER
QAB ERK EKTE XO BXSXBKB XAEZ
AND THE TEXT IS DIVIDED INTO
V G WIZUHO ZC CXSK FKEEKIO EZ
GROUPS OF FIVE LETTERS TO
XAOUIK QMMUIQMW XA EKFKVIQHRXM
INSURE ACCURACY IN TELEGRAPHIC
EIQAOYXOOXZA
TRANSMISSION

PROBLEM No.5

Naval Text

FENFWEN DH AND OXZNEVWP WD
PREPARE TO GET UNDERWAY AT
SNEH NBAUD UOXZENZ FROM TBLN
ZERO EIGHT HUNDRED PLUS FIVE
DBJN THOEDNNX WFEBR FNEBHZ IN
TIME FOURTEEN APRIL PERIOD BE
FENFWENZ THE JWGBJCJ MFNNZ
PREPARED FOR MAXIMUM SPEED
DVNXDP YXHDM
TWELVE KNOTS

PROBLEM No.6

Non-Naval Text

N IXTRYILTNO NOYWNZXC RF LAX
A RECIPROCAL ALPHABET IS ONE
RA HWRTW NOO CWX KNOMXF RA
IN WHICH ALL THE VALUES IN
CWX NOYWNZXC NIX IXTRYILTNO
THE ALPHABET ARE RECIPROCAL
RA YNRIF
IN PAIRS

PROBLEM No.7

```
DO  IBWY  UTXB  CGESWBS  SBVOWYRBW
AT  ZERO  FIVE  HUNDRED  DESTROYER
STXTVTYE  VBXBEOBBE  ZCDEMB  ZYGWVB
DIVISION  SEVENTEEN  CHANGE  COURSE
OY  OCWBB  UTXB  UYGW
TO  THREE  FIVE  FOUR
```

Note: - The word "destrcyer" is believed to be in the text of this message.

PROBLEM No.8

```
ALTDM  VSRJBXTY  LIPRN  AVJU
       E           O     E
ISBDC  FNWOSBQRRV  AK  SVCCOTM
                L       E
VDY  LPUMA  QRBPFIO  JK  NP WFDLDEO
E     C        O     T    O
TVQH
 E
```

If solution is not achieved in 45 minutes, break the seal and read the next page.

SUBMITTED BY ENSIGN [illegible].EWING,U.S.N.,U.S.S.HERBERT

Problem No. 8 cannot be solved. It is a meaningless jumble of letters written at random.

Postmortem

Having solved problems No.1 to 7, and having learned that No.8 is really not a cryptogram but a hodgepodge of letters, the student should be impressed with the fact that problems Nos. 1 to 7 can be solved because language is hidden by the cipher and No.8 cannot be solved because there is no language there. Furthermore, he has seen how simple this type of problem becomes when there is a "Known" word as in Problem No.7

NAVY DEPARTMENT
Office of Chief of Naval Operations
WASHINGTON

Op-20-GYT

~~RESTRICTED~~

ELEMENTARY COURSE IN CRYPTANALYSIS

- ASSIGNMENT No.2 -

****

PART I

1. The problems in Assignment No. 1 were solved by inspection, illustrating the fundamental principle that cipher solution is based on the peculiarities of the underlying plain text. Words were assumed instead of individual letters, which led to rapid solution of the cryptograms. Before proceeding to more complex types of ciphers, a brief description of the individual letter distribution is given.

The English language is written by means of 26 characters called letters, which, taken together and considered as a sequence of characters, constitute an alphabet. Nearly all written languages are similar, but there are a few exceptions, notably Chinese. The principles discussed herein concerning the characteristics of English apply to all modern languages of alphabetical construction.

2. If a tabulation of the occurrence of individual letters, called a frequency table, is made of a large volume of ordinary Naval text (nearly but not quite identical with English literary text), some interesting facts are disclosed. The Mechanics of English Table shows graphically the relative frequency of each individual letter to be expected in 200 letters of Naval Text (based on an actual count of 20,000 letters of text). Note that the most frequent letters are E,T,O,N,A,I,R, and S, and the most infrequent are J,K,Q,X, and Z.

Just as single letters have characteristic frequencies, pair of letters, called digraphs, and sets of three letters, called trigraphs, do also. These tables are also given under Mechanics of English.

3. Frequency tables should be used only as a check on assumptions. A very common fault among amateur cryptanalysts is the placing of too much weight on the frequencies of individual letters. For instance, "E" and "T" have the two highest average values in English text, but they are not necessarily the highest-frequency letters in a given cryptogram. Repetitions and peculiar letter distributions are far more important than frequencies. As an example of the above principles, a 4-letter repetition is found in the text and there is strong evidence to show that these 4 letters are word endings. Since it is a repetition, it probably is a common word ending. If no previous correct assumptions had been made, the decision between the common endings - ENCE, MENT, TION, and ING must be made. Here the frequency table comes into play for the first time. All of the letters involved are high frequency letters excepting M, C and G. M occurs as the 1st letter of the repetition. C occurs as the 3rd letter and G as the 4th. The frequency table usually is very helpful in choosing the correct possibility, but even in such a case it cannot be relied upon completely. With limited text, or text containing unusual language, frequency tables must be viewed with suspicion.

4. Another application of the frequency table is its use in identifying vowels and high-frequency consonants. With limited text, repetitions may not occur, or the cipher system may be sufficiently complex to conceal repetitions in the plain text. As a measure which is more or less a last resort, vowels are classified as such, not individually as A, E, etc., but as a class. Before attempting this, a study of the digraphic frequency table shows that in general vowels combine infrequently with vowels, but they do combine frequently with both high and low frequency consonants; that high frequency consonants combine most frequently with vowels and other high frequency consonants; and that low frequency consonants combine most frequently with vowels. Vowel classification in a complicated system leads up to the point where "assumptions that fit" can be made. Even here the frequency table is only a guide, and sometimes an unreliable guide.

5. Recently (March, 1937) an author published a book of over 50,000 words in which the letter "E" does not appear at all. The book is readable and the sentences are not jerky or awkward. In normal English the six vowels A, E, I, O, U, Y represent 40% of the total text. Of these, the value of E alone is 13%. Yet in a book of large volume without a single "E", the percentage of vowels used still must closely approximate the same value, 40%. That is, the number of vowels as a class, can still be depended upon and if E does not appear, the other vowels will be used with greater than normal frequency to compensate for its omission.

6. Just as vowels represent a definite percentage of the entire text, the low frequency consonants J, K, Q, X, Z, together represent a definite percentage of less than 2%. One or more of these letters may vary considerably from its normal frequency in a given amount of text, but the percentage of the group will remain less than 2%.

7. Another use of the frequency table involves the classification of both vowels and consonants. In vowel classification it is usually possible to classify as vowels the letters representing A, E, I, and O without difficulty, but U and Y are almost impossible to identify as vowels. Therefore, in connection with vowel classification, the classification of groups as high, intermediate, and low frequency is helpful. The eight high frequency letters E, T, O, A, N, I, R, S comprise 66½% of the text. Of this amount, the four vowels E, O, A, and I are 36½% and the consonants T, N, R, and S, 30%. The other 18 letters, including the low frequency group J, K, Q, X, Z, comprise the other 1/3 of the text. It is usually easy to pick the 8 high frequency letters of the cipher text with reasonable assurance that they represent at least 7 of the 8 high frequency letters of English because, as the frequency table shows, the values of the next highest frequency letters after S drop sharply. Of the 8 highest frequency letters, it is possible to classify 4 vowels, as explained previously, leaving the other 4 automatically classed as being in the T, N, R, S group. Thus with 4 vowels, 4 high frequency consonants, and 5 low frequency letters classified, the problem of making correct assumptions to fit the cipher text is simplified.

8. The foregoing discussion has been concerned only with the English language. English is one of the most difficult of languages for the cryptanalyst. French and German, for example, both show E as outstandingly high, much more so than in English, and this letter can be spotted at once from the frequency table of the proper alphabet. Also, these languages have certain invariable high frequency combinations such as the German CH and the French or Spanich QU, which aid analysis to a great degree. Such language characteristics undoubtedly have led European authors of works on this subject to stress the value of individual letter frequencies far beyond the point where they can be depended upon.

9. In all but the simplest problems, a frequency table is constructed for use as a guide, as explained in the foregoing paragraphs. To construct a frequency table, the A's, B's, etc., of the cryptogram are counted. It is usually best to do this

graphically, as shown in the Mechanics of English Table. The reason for this will become apparent in later assignments. It is also beneficial to make a Trigraphic Frequency Table. This is done by listing, for each letter of the alphabet, A, for example, the letter which precedes (prefix) and the letter which follows (suffix) for each appearance of A in the text. For the following cipher text - B A D V B C A Q R B A D L P R A S W B Q A, a partial (for A and B only) trigraphic table is:

```
     B C B R Q
  A
     D Q D S -
    ___________
     - V R W
  B
     A C A Q
```

The upper line of letters listed with A represents the prefix in their order of occurrence, the lower line gives the corresponding suffixes. This table shows at a glance the digraphs, trigraphs, and repetitions in the message. It is the only sure way of locating all repetitions in a long cryptogram, and it is valuable in classifying vowels.

10. In the Mechanics of English table, the frequency of initial and final letters is also given. This should be used in the same manner as any other frequency table: merely an aid and not a sign post.

MECHANICS OF ENGLISH TABLE (For Naval Text)

Frequency of Individual Letters to be expected in 200 letters of Naval Text. (Based on a count of 20,000 letters).

15	A	~~////~~ ~~////~~ ~~////~~
3	B	///
6	C	~~////~~ /
9	D	~~////~~ ////
26	E	~~////~~ ~~////~~ ~~////~~ ~~////~~ ~~////~~ /
5	F	~~////~~
5	G	~~////~~
5	H	~~////~~
15	I	~~////~~ ~~////~~ ~~////~~
	J	
1	K	/
6	L	~~////~~ /
4	M	////
16	N	~~////~~ ~~////~~ ~~////~~ /
17	O	~~////~~ ~~////~~ ~~////~~ //
5	P	~~////~~
	Q	
15	R	~~////~~ ~~////~~ ~~////~~
11	S	~~////~~ ~~////~~ /
18	T	~~////~~ ~~////~~ ~~////~~ ///
6	U	~~////~~ /
3	V	///
3	W	///
1	X	/
3	Y	///
1	Z	/

Frequency of Digraphs and Trigraphs to be expected in 2,000 letters of Naval text. (Based on a count of 20,000 letters)

Most Frequent Digraphs

ER-43	RO-24	OR-20
IN-42	ES-23	OU-20
ON-38	ST-23	RI-19
EN-34	TI-23	ET-18
RE-34	CO-22	PE-18
AT-31	ND-22	VE-17
AN-29	NE-22	AR-16
NT-27	NG-22	TA-16
TE-27	TO-22	DE-15
EE-25	IO-21	LE-15
ED-24	TH-21	SE-15

Most Frequent Trigraphs

ING-17	ERI-9	ATT-6
ENT-13	ION-9	DRE-6
ERO-11	PER-9	LAN-6
EEN-10	TEE-9	ONE-6
GHT-10	COU-8	RED-6
IGH-10	IVE-8	RIN-6
TIO-10	OUR-8	RIO-6
ZER-10	OUT-8	TER-6
AND-9	EST-7	TIN-6
	ATI-6	

FREQUENCY OF INITIAL AND FINAL LETTERS

Letters -	A	B	C	E	D	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Initial -	9	6	6	5	2	4	2	3	3	1	1	2	4	2	10	2	-	4	5	17	2	-	7	-	3	-
Final -	1	-	-	10	17	6	4	2	-	-	1	6	1	9	4	1	-	8	9	11	1	-	1	-	8	-

~~RESTRICTED~~

FIRST LETTER

SECOND LETTER	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	1	4	7	10	9	2	4	6	2	-	1	14	11	10	3	4	-	14	8	16	4	3	4	-	3	-
B	4	-	-	2	3	-	-	-	1	-	1	1	1	3	3	-	-	2	2	1	3	-	-	-	1	-
C	9	-	1	2	8	1	1	-	6	-	1	1	-	8	3	-	-	6	8	4	1	-	-	1	2	-
D	5	1	-	2	24	-	1	-	2	-	-	-	-	22	11	-	-	9	1	3	2	-	-	-	1	-
E	-	8	10	15	25	2	8	7	2	-	3	15	5	22	4	18	-	34	15	27	3	17	10	-	2	10
F	2	-	-	3	7	1	1	1	5	-	1	1	-	5	9	-	-	2	2	4	-	-	-	1	3	-
G	3	-	2	1	1	-	1	-	11	-	-	-	-	22	2	-	-	3	-	1	1	-	-	-	-	-
H	1	-	5	1	2	-	11	-	-	-	-	-	-	2	-	1	-	-	4	21	-	-	1	1	-	-
I	8	1	1	12	7	13	2	6	-	-	2	8	4	6	6	1	-	19	13	23	4	4	5	4	1	-
J	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
K	3	-	4	-	1	-	-	-	-	-	-	-	-	2	1	-	-	1	-	-	-	-	-	-	-	-
L	7	6	-	2	6	5	2		8	-	1	6	1	2	4	5	-	1	2	3	2	-	-	1	3	-
M	2	2	-	2	6	1	-	1	2	-	-	-	2	2	9	-	-	3	1	1	2	-	-	-	2	-
N	29	-	-	-	34	-	1	-	42	-	1	-	-	3	38	-	-	3	1	2	9	-	-	-	-	-
O	-	4	22	4	6	12	2	3	21	2	-	7	4	10	2	7	-	24	5	22	-	1	6	1	2	-
P	4	-	1	3	10	1	1	1	2	-	-	2	2	2	8	3	-	2	6	3	1	-	-	-	2	-
Q	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-
R	16	-	4	8	43	2	2	7	10	-	-	1	-	2	20	8	-	2	1	10	1	-	1	-	1	-
S	11	-	1	6	23	1	2	1	10	-	-	2	1	9	9	3	-	8	6	8	4	-	-	-	2	-
T	31	-	4	6	18	5	6	11	11	-	-	1	-	27	7	2	-	10	23	8	10	-	-	2	2	-
U	1	1	-	2	1	1	3	6	-	-	-	1	-	3	20	1	3	4	6	4	-	-	-	-	-	-
V	3	-	-	1	7	1	1	1	9	-	-	2	-	-	1	-	-	-	-	-	-	-	-	-	-	1
W	-	-	-	1	2	-	-	1	-	-	-	-	-	1	4	-	-	1	3	12	-	-	-	-	1	-
X	1	-	-	-	4	-	-	-	5	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
Y	5	2	1	3	1	1	-	-	-	-	-	2	3	-	1	-	-	-	-	8	-	-	-	-	-	-
Z	-	-	-	-	4	-	1	-	-	-	-	-	-	-	1	-	-	-	-	4	-	-	-	-	-	-
	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

Digraphs to be expected in 2,000 letters of Naval text. (Based on a count of 20,000 letters).

- PART II -

SYSTEMATICALLY MIXED CIPHER ALPHABETS

1. Any system which will permit the derivation of a sequence of letters from an easily memorized key, may be used to construct a systematically-mixed cipher alphabet. One of the most useful types is the keyword-mixed sequence. In this type the keyword or keyphrase is written down, repeated letters, if any, being omitted after their first occurrence; then the remaining letters of the alphabet are written in their normal order, omitting such letters as already occur in the key.

Example -

Let the keyword be WASHINGTON. The corresponding mixed sequence becomes:

WASHINGTOBCDEFJKLMPQRUVXYZ

2. Although transposition methods have not yet been discussed, it will be necessary to demonstrate how these may be applied to keyword-mixed sequences to further disarrange the sequence.

Example -

Three examples will be given using the keyword RENDEZVOUS. The keyword-mixed sequence may be written:

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

and the columns taken off so as to form the following sequence:

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

The alternate columns may be reversed to obtain this sequence.

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

Also, a numerical key, derived from the keyword itself, may be applied to vary the other in which the columns are taken off:

```
5-2-3-1-9-8-4-7-6
R E N D Z V O U S
A B C F G H I J K
L M P Q T W X Y
```

The transposition-mixed sequence now becomes:

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

3. Once aware of such systems of constructing cipher alphabets, it is comparatively easy to rebuild the generating figure. Note that, in example (1), W, X, and Y are three letters apart with H, I, and J to their left, respectively. This suggests that W, X, and Y are on the bottom line of the generating figure, H, I and J on the next line above, and that V, C, and U are in the keyword.

In example (2), the presence of LM, PQ, and XY in their normally adjacent positions suggests that the alternate columns have been reversed, which is checked by the A and B on either side of LM, and the I and J on either side of XY

In example (3), note again the HW, IX, and JY combinations which suggest a columnar system and may be used to rebuild the original sigure in much the same manner as in the case of the simple columnar transposition.

4. Another simple method of producing a systematically-mixed alphabet is called the decimation method. The basic sequence to be decimated is regarded as a circle, and the letters are counted off and written down in a separate list. When a letter has been used in the final sequence, it is eliminated from the basic sequence before the process continues.

Example -

Suppose the number agreed upon is 7, and the basic sequence to be decimated is a normal alphabet. The letters will be taken from the basic sequence, after counting off, in the following order:

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
16	4	10	20	19	21	1	18	8	5	13	15	11	2	24	22	17	6	9	25	3	26	14	12	23	7

Mixed Sequence

GNUBJRZISCMXKWLAQHEDFPYOTV

5. Almost any transposition method may be applied in the construction of systematically-mixed cipher alphabets. Practical considerations limit the complexities which may be introduced, and the greatest amount of mixing by systematic processes will give no more security than that resulting from a random selection.

6. During the process of solution of any cryptogram, much labor can often be avoided by a reconstruction of the system used, when only a portion of the simpler types have been recovered. In any case, the solution of a cryptogram should never be considered complete until the system used has been determined and reconstructed, insofar as the available material permits.

PRIMARY AND SECONDARY ALPHABETS

7. It is obvious that the cipher component of a cipher alphabet may be shifted or slid against the plain component at 26 points of coincidence so as to produce a series of different enciphering alphabets. The primary alphabet is the basic arrangement of the original sequences, and the derived alphabets are called secondary alphabets.

8. In producing the secondary alphabets the primary alphabet may be arranged as follows:

(a) The same sequence may be used as both the plain and cipher components, and slid against itself.

Example -

```
WASHINGTOBCDEFJKLMPQRUVXYZWASHINGTOBCDEFJKL
   WASHINGTOBCDEFJKLMPQRUVXYZ
```

(b) The primary cipher component may be slid against the normal sequence.

Example -

```
ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQ
     WASHINGTOBCDEFJKLMPQRUVXYZ
```

(c) The primary plain and cipher components may be different mixed sequences.

Example -

```
GOVERNMTABCDFHIJKLPQSUWXYZGOVERNMTABCDFHIJK
    WASHINGTOBCDEFJKLMPQRUVXYZ
```

9. When the plain component of a cipher alphabet is a normal sequence, as in par. 8(b) above, the original cipher sequence becomes evident as soon as the enciphering alphabet is reconstructed. However, when the enciphering alphabets of the type described in par. 8(a) and (c) are reconstructed with the plain components in normal order, the original sequences are not apparent.

10. The cipher alphabet in par. 8(a) would appear as follows when obtained after the solution of a cryptogram employing this alphabet:

Plain - ABCDEFGHIJKLMNOPQRSTUVWXYZ

Cipher - YGTOBCHWADEFJSNKLMZIPQXRUV

Note the letters underlined, which indicate by their normally adjacent positions that these letters are adjacent in the primary cipher alphabet, which is reconstructed as follows:

Plain - EF JKL M PQR UV X YZ WASHINGTOBCD

Cipher - BC DEF J KLM PQ R UV XYZWASHINGTO

The letters not underlined are fitted in their proper locations, which are assumed from a knowledge of the possible constructions of the original sequence.

11. The cipher alphabet in par. 8(c) would appear as follows when the plain component is in normal order:

Plain - ABCDEFGHIJKLMNOPQRSTUVWXYZ

Cipher - INGTZOVBCDEFSAXJKWLHMYPQRU

This alphabet may be rearranged in its original sequence in much the same manner as illustrated in the preceeding paragraph:

SOLUTION BY COMPLETING THE PLAIN COMPONENT

12. This is a very useful and rapid mechanical method of solving cryptograms when both the plain and cipher components are known sequences, but when their point or points of coincidence are unknown.

13. Consider the problem in which a direct standard cipher alphabet has been used. If we complete the normal alphabet sequence in a column under each cipher letter, the result is the same as having tried the cipher component in each of the 25 possible points of coincidence with the plain component, and having applied the resulting deciphering alphabets to the cipher text.

14. If the first ten letters of the cipher text are FTUEETADF, the solution by completing the plain component will appear as follows:

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

An examination of the successive horizontal lines, called generatrices, (singular generatrix), discloses one and only one line of plain text: THIS SHORT. Instead of laboriously writing down the several columns, it is recommendeed that the student prepare a set of alphabet strips, each repeated so that every strip will contain 52 letters, and mount them upon some rigid material convenient to handle. Such a set of sliding alphabets will be found exceedingly valuable in all work of this kind.

15. Next consider the problem in which the cipher alphabet employed is any type other than a direct standard cipher alphabet.

16. In this case an additional step is necessary before completing the plain component sequence. In order to obtain the same result as having applied each of the 26 deciphering alphabets to the cipher text, the cipher letters must first be converted into their plain component equivalents. To find the plain component equivalents the cipher alphabet is written with both components in their original order, and placed at any point of coincidence.

17. Let us suppose the following random mixed cipher alphabet has been recovered from the solution of earlier cryptograms:

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

Also suppose another cryptogram, which begins FENFWENDHA, etc., is suspected of employing one of the secondary alphabets derived from this primary alphabet, that is, the same system has been used with a different key.

First convert the cipher letters into their plain component equivalents. Then use the normal alphabet sliding strips to complete the normal alphabet sequence beneath each plain component equivalent.

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

This example will demonstrate that although the whole series of values may be changed by merely shifting the cipher component to another point of coincidence, the solution of a cryptogram in a different key was obtained very easily, without any frequency table analysis.

Had the plain component been a mixed sequence also, the solution would proceed as in this example except that the original plain component sequence would be used in completing the sequence beneath each plain component equivalent, instead of the normal alphabet strips.

ELEMENTARY COURSE IN CRYPTANALYSIS

- ASSIGNMENT No. 2 -

PART III -

1. Define the word generatrix A line or column of cipher text separated from a similar line of column of plain text by a certain interval.

2. Answer the following questions:

   (a) What characteristics of a systematically-mixed keyword alphabet aid in the recovery of the keyword? Certain low frequency letters usually not in the keyword give away the "system" by their arrangement

   (b) What is meant by converting the cipher text into its plain component equivalents? If the system has been recovered and a secondary alphabet used, the first step is to decipher as if the primary alphabets were used.

   (c) What types of mono-alphabet substitution ciphers can be solved by the use of sliding strips alone? Ciphers in the same system with secondary alphabets

   (d) Solve the attached problems and reconstruct the systems used.

Submitted by Ensign Ronald K. Irving USN

U.S.S. Herbert

Navy Yard

Philadelphia, Pa.

- PROBLEMS TO ASSIGNMENT No. 2 -

ABCDEFGHIJKLMNOPQRSTUVWXYZ
WESTRNUIOLGAPHCMYBDFJKQVXZ

PROBLEM No. 1

Naval Text

Keyword - Western Union Telegraph Company

KODJWA SCHFWSF HCF XRF PWTR
VISUAL CONTACT NOT YET MADE
QOFI TRDFBCXRB TRFWSIPRHF
WITH DESTROYER DETACHMENT
MRBOCT WBUCHHR TOBRSFRT BRPWOH
PERIOD ARGONNE DIRECTED REMAIN
KOSOHOFX BRHTRZKCJD FC PWGR
VICINITY RENDEZVOUS TO MAKE
SCHFWSF WHT TRAOKRB SCMORD
CONTACT AND DELIVER COPIES
WAA DRSBRF WHT SCHNOTRHFOWA
ALL SECRET AND CONFIDENTIAL
MJEAOSWFOCHD
PUBLICATIONS

PROBLEM No. 2

January February March Keyword

ZNRD BUTDUWDG WMD HBZXW
ROME INVENTED THE FIRST
XVKXWEWVWBNU XPXWDR FUNSU
SUBSTITUTION SYSTEM KNOWN
EVYBVX ILDXLZ VXDG LU LYAMLKDW
JULIUS CAESAR USED AN ALPHABET
BU SMBIM G SLX XVKXWBWVWDG HNZ
IN WHICH D WAS SUBSTITUTED FOR
L INRRIL D HNZ K INRRIL H HNZ I
A COMMA E FOR B COMMA F FOR C
DWI BU MBX INZZDXANUGDUID
ETC IN HIS CORRESPONDENCE

PROBLEM No. 3

Keyword - Black

WUBGT BGGJKUP MBMVJML ZJZSUUM
PEEKI ARRIVIS NANKING FIFTEEN
EXMPGUP SEJN PBSU ZRGUJLM FUM
HUNDRED THIS DATE FOREIGN MEN
RZ HBG GUNUMS SHR DGJSJNE
OF WAR PRESENT TWO BRITISH
LXMDRBSN RMU QBWBMUNU PUNSGRTUG
GUNBOATS ONE JAPANESE DESTROYER
BS WGUNUMS NR BWWGUEUMNJRM
AT PRESENT NO APPREHENSION
BFRML ZRGUJLMUGN IJST CXJUS
AMONG FOREIGNERS CITY QUIET

PROBLEM No. 4 Keyword - Cryptograph Non-Naval Text

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

Standard cipher alphabet with numerical

PROBLEM No.5 key - 7-1-3-6-2-5-4-8 resulting BJRZEMUCKSGOWFNVD etc. Non-Naval Text

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

PROBLEM No. 6 Keyword - Operations

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

~~RESTRICTED~~

OP-20-G

NAVY DEPARTMENT
Office of Chief of Naval Operations
WASHINGTON

*

ELEMENTARY COURSE IN CRYPTANALYSIS

- ASSIGNMENT No.3 -

Paragraph

NUMERICAL CIPHER ALPHABETS

1. Cipher alphabets whose cipher components consist of numbers are practicable for telegraph or radio transmission. They may take forms corresponding with those employing letters.

(a) Standard numerical cipher alphabets are those in which the cipher component is a normal sequence of numbers.

Example -

Standard numerical cipher alphabet

Plain -	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Cipher -	11	12	13	14	15	16	17	18	19	20	21	22	23	24

O	P	Q	R	S	T	U	V	W	X	Y	Z
25	26	27	28	29	30	31	32	33	34	35	36

095

Since there are but ten digits, it is obvious that, in order to represent a complete alphabet, combinations of at least two digits are necessary.

(b) Mixed numerical cipher alphabets are those in which the cipher component is not a normal sequence of numbers.

Example -

(1) Random mixed numerical cipher alphabet

Plain -	A	B	C	D	E	F	G	H	I-J	K	L	M	N	O
Cipher -	76	88	1	67	4	89	66	99	96	2	69	90	77	5

P	Q	R	S	T	U	V	W	X	Y	Z
87	60	39	79	3	78	68	98	86	70	97

This example will also illustrate a type of numerical cipher alphabet in which some of the digits may be employed singly and some in pairs to represent single plain-text letters, thus retarding the attempts of cryptanalysts to isolate the individual cipher equivalents of plain-text letters after they have been run together in the cryptogram.

(2) Systematically mixed numerical cipher alphabet

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

The pair of numbers which appear as row and column indicators are used as the cipher equivalent of the plain letter found at the intersection of the row and column. That is, A plain is 11 cipher, B plain is 12 cipher, etc.

Rectangles of various shapes and sizes may be used, having various key number arrangements, and including cells for proper names and places or blank cells. Also, the plain alphabet may be any type of mixed alphabet, and may be inscribed by following any prearranged route to fill the proper cells of the rectangle.

2. Numerical cipher values lend themselves to treatment by various mathematical processes to further complicate the cipher system in which they are used. These processes, usually addition or subtraction, may be applied to each cipher equivalent individually, or to the complete numerical cipher message by considering it as one number.

CIPHER ALPHABETS EMPLOYING VARIANTS

3. In order to disguise, suppress, or eliminate the characteristic frequencies of the plain-text letters, cipher alphabets may be made up with variant values in their cipher components.

4. An equal number of cipher values may be assigned each plain-text letter, usually by means of a systematic arrangement, or a set of values may be assigned each plain-text letter in accordance with its relative frequency in ordinary plain language.

(RETYPED FOR PURPOSE OF CLARITY)

This example will also illustrate a type of numerical cipher alphabet in which some of the digits may be employed singly and some in pairs to represent single plain-text letters, thus retarding the attempts of cryptanalysts to isolate the individual cipher equivalents of plain-text letters after they have been run together in the cryptogram.

(2) Systematically mixed numerical cipher alphabet

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

The pair of numbers which appear as row and column indicators are used as the cipher equivalent of the plain letter found at the intersection of the row and column. That is, A plain is 11 cipher, B plain is 12 cipher, etc.

Rectangles of various shapes and sizes may be used, having various key number arrangements, and including cells for proper names and places or blank cells, Also, the plain alphabet may be any type of mixed alphabet, and may be inscribed by following any prearranged route to fill the proper cells of the rectangle.

2. Numerical cipher values lent themselves to treatment by various mathematical processes to further complicate the cipher system in which they are used. These processes, usually addition or subtraction, may be applied to each cipher equivalent individually, or to the complete numerical cipher message by considering it as one number.

CIPHER ALPHABETS EMPLOYING VARIANTS

3. In order to disguise, suppress, or eliminate the characteristic frequencies of the plain-text letters, cipher alphabets may be made up with variant values in their cipher components.

4. An equal number of cipher values may be assigned each plain-text letter, usually by means of a systematic arrangement, or a set of values may be assigned each plain-text letter in accordance with its relative frequency in ordinary plain language.

5. A system which provides twelve variants of letter pairs for each plain letter may be constructed as follows:

Let the keywords be BALTIMORE and MARYLAND. The corresponding keyword alphabets become:

(1) BALTIMORECDFGHJKNPQSUVWXYZ, and

(2) MARYLNDBCEFGHIJKOPQSTUVWXZ

The letters of the first keyword sequence are used as the row and column indicators of a 25 cell rectangle, and those of the second keyword sequence are inscribed within the cells of the rectangle according to a diagonal route.

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

In this example, A plain may be represented by any one of the following cipher equivalents:

BN, BV, MN, MV, DN, DV,

NB, NM, ND, VB, VM, or VD.

6. There are not only numerous variations in the use of rectangles, but many types of lists and tables may be employed in the construction of miscellaneous types of cipher alphabets. The practical disadvantages in the use of most of these miscellaneous types in mono-alphabetic substitution are not compensated by any real gain in cryptographic security.

NOTES ON PREPARATION OF WORK SHEETS

7. Cross-section paper with one quarter inch squares makes the best work-sheet. A typewritten work-sheet is nearly as good, for it is the even spacing which is essential. Three spaces should be left between lines so as not to overcrowd the work-sheet. Use printed block capital letters. Colored pencils are helpful in marking off repetitions and peculiarities of letter distribution.

8. Over each cipher equivalent of a plain-text value write its frequency. Underscore all repetitions and reversible digraphs. Examine the text and overscore any peculiairites of letter distribution. Recording the frequencies on the work-sheet is of the greatest importance when dealing with a minimum of text. It saves constant reference to the frequency tables, which interrupts the train of thought. It saves considerable time in the end, and might mean the difference between success and failure in a complex problem.

OUTLINE OF CIPHER SOLUTION

9. The solution of a substitution cipher generally progresses through the following stages:

(a) Analysis of the cryptogram(s).
  (1) Preparation of frequency table.
  (2) Search for repetitions.
  (3) Determination of the type of system used.
  (4) Preparation of work sheet.
  (5) Preparation of frequency tables for the individual cipher alphabets (if more than one).
  (6) Tabulation of long repetitions and peculiar letter distributions.
(b) Classification of vowels and consonants by a study of:
  (1) Frequencies
  (2) Spacing
  (3) Letter combinations.
  (4) Repetitions.
(c) Identification of letters.
  (1) "Breaking in" process.
  (2) Verification of assumptions.
  (3) Filling in good values throughout messages.
  (4) Recovery of new values to complete the solution.
(d) Reconstruction of system.
  (1) Rebuilding of the enciphering table.
  (2) Recovery of key(s) used in the operation of the system.
  (3) Recovery of the key or keyword(s) used to construct the alphabet sequences.

10. No outline can be made to suit all cipher solutions, because special conditions may call for short cuts or extra steps in solving a particular problem. Cipher solution is by no means an exact mechanical process, however the object of giving an outline is to show that success in cipher solution is the result of orderly reasoning.

11. Determination of the type of cipher system used in a given cryptogram is often the most difficult step in cryptanalysis. The student should notice the external characteristics of each new type studied, because a comparison of these characteristics is the basis for determining the type of system used in an unsolved cryptogram.

PRINCIPLES INVOLVED IN CIPHER SOLUTION

12. Whenever possible, classify the vowels and consonants before assuming values. The four considerations in distinguishing the vowels from the consonants are as follows:

(a) The low frequency values are almost invariably consonants of low or medium frequency. The intermediate frequency values are usually consonants but may be vowels. They cannot be classified except as they combine with letters already classified and are the most difficult to classify. The high frequency values are either the vowels "A,E,I,O" or consonants of high frequency.

(b) It is unusual to find over two or three consonants of low frequency in combination. Vowels usually stand alone - combinations of more than two vowels are extremely rare. A gap of six or eight letters between two known vowels indicates the need of one or more intermediate vowels.

(c) Consonants combine with vowels, most of which are of high frequency. Vowels combine with consonants, many of which are of low frequency. Letters associated with low frequency values are vowels. Letters associated with high frequency values are consonants.

(d) Of the 30 most frequent letter pairs, 22 are vowel-consonant or consonant-vowel, 5 are consonant-consonant, and 3 are vowel-vowel combinations. Repetitions in the cipher text indicate high frequency letter combinations. Therefore, the repetitions of a given letter combination creates the presumption that one of the letters is a vowel and the other a consonant.

"U" is of low frequency and can be classified only by "spacing" after A,E,I and O have been classified. The vowel of 5th highest frequency in an alphabet is almost invariably a "U". It is usually impossible to classify "Y" as a vowel - partly on account of its very low frequency and partly because "Y" is sometimes a consonant.

Mark each vowel by a circle as soon as classified - both on the work sheet and the frequency table. Values identified as consonants should be marked by an overscore or some similar method.

13. The frequency table is only a guide in the identification of letters, and sometimes an unreliable guide. Repetitions are far more important than frequencies in the identification of letters. "E" is one of the poorest letters to identify first, as it combines with so many letters

that it does not help in further identifications. "E" will always be discovered without special search. "N" is probably the most valuable letter to identify first, (and one of the easiest) on account of its frequent occurrence in "ING", "ENT", "AND", and "ION". Do not disregard the low frequency letters. A "G" may disclose an "N" or a "Q" show the "U" following it.

14. Do not force the solution by attempting to make a logical assumption prove correct when it cannot be verified. The attack should always follow the line of least resistance. Find a weak point in the cryptogram and then work on it until the cipher is broken. The beginning and end of a message are always weak, and there are usually several other good points of attack.

15. Do not give up an assumption too easily, but do not cling to it too long. Experience is the only teacher as to the time which should be spent on a given assumption. Consider what words would probably or even could possibly appear in the cryptogram, then try to fit them in. Check the letter values of the assumed words in a few places before filling in the assumed values throughout the cryptogram.

16. As far as possible, assume words or phrases with one or more letters repeated in them. Then fit them to the cipher text where the same peciliarities of letter distribution are found.

Example. -

LVKKVKKVNNV	XNOAVJRJKOBDB	XFXHS
MISSISSIPPI	CRYPTANALYSIS	ENEMY

When repeated letters cannot be used to fit a word to the cipher text, the frequencies of the letters and the location of the vowels are nearly as good peculiarities of letter distribution on which to base an assumption.

17. In 1841, Edgar Allen Poe made the following significant statement which still remains of interest to present day students of cryptanalysis:

> "The basis of the whole art of cipher solution is found in the general principles of the formation of language itself, and is thus altogether independent of the particular laws which govern any cipher, or the construction of its key".

ELEMENTARY
COURSE IN CRYPTANALYSIS

- ASSIGNMENT No.3 -

PART II - Home Work

Solve the following cryptograms. Naval telegraphic text has been used to give a certain degree of familiarity with naval language and to aid the student in making assumptions.

The same general technique used in solving the problems of the first two assignments will also assure solution of these problems.

Reconstruct the systems used in each problem.

- 1 -

```
06021   00501   01051   52202   06082
 R  E
32510   08040   22109   08040   82211

08041   71513   14222   10224   02012

20202   01081   90615   17080   11122

14020   11906   05100   20211   22140

62319   05150   12213   02050   61302

05011   00523   06210   22214   06020

22214   06020   22602   06052   11902

02112   20302   17240   21902   06150

51106   02190   50622   01050   50119

05211   62215   05012   20518   05060

60505
```

Solution: RECONNOITER AUX CAYES BAY AT DAYLIGHT SEVENTEEN APRIL AND THEN PROCEED THRU POINT GEORGE ON COURSE THREE THREE ZERO SPEED TWELVE PERIOD REPORT NOON POSITION TOMORROW.

Keyword - New York (cipher component arranged numerically from 01 to 26)

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

Station:—

Weather forecast Thursday partly cloudy with scattered showers easterly winds at surface averaging twenty five knots visibility less than average undesirable flying condition and height of low clouds at present about one thousand feet

Keyword: monday   Key arrangement:—

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

```
ACUEI   AIOIA   UEEUE   UAIIA   IOIAU
EEAAO   UEUEA   OEIIU   EUEOE   EEAIA
IOAUE   UOAEU   UOIOA   IIEUE   OIAII
IEUAO   AUEAE   EOIOE   EEIOA   IIEOO
OAAOA   IEUAE   AAAIE   OEUUA   AOAIU
EIOAO   IOUAE   IUAAO   IAUEI   AIOUA
EUUEI   IUAEU   EUOEI   OAIAA   UEEUA
EOOOA   AOAIE   UAEEE   OIOEE   EIOAI
IEOOO   IAEIO   UEIOO   IAIAO   OAUEI
OUEAA   EOEUO   IAIEU   IAOIA   AUAUE
IIOIA   AAEAO   AIEOI   AEUEU   UEAIO
IEOAO   OEAOI   EEUUE   OIAEA   OUAAO
UEOII   EAO
```

Solution :-
At fifteen fifty attacked by [illegible] Lexington near [illegible]
bombing & [illegible] [illegible] [illegible] three hundred [illegible]
bombing position altitude one [illegible] thousand feet no damage
to Saratoga
Signed [illegible] U.S. Navy
Key arrangement :-

	A	E	I	O	U
A	D	Y	[illegible]	[illegible]	R
E	S	B	[illegible]	M	T
I	[illegible]	C	[illegible]	O	W
O	F	P	[illegible]	P	X
U	V	E	K	G	Z

```
IDUGE   JFOIK   IQPED   IALUM   WIECZ
AGUHA   QIVEC   EIKUG   KILAE   PQIKI
FOGUA   ZIKOP   EPIQE   JAZQI   QIDIK
IBAIQ   HAIKP   OOLAH   IQWIU   GAHAL
KIQII   DEVOL   AHJEG   UIKLA   PEIQO
LAHOI   OFALH   APEZA   QIEPO   RLOZA
IKOPK   IBALO   EPQIJ   EAZIQ   QIDIL
DGUIN   IKIQQ   IHAWI   UGHAL   AKIQI
IDCEI   KHAUG   AHQII   KQIIQ   HAROO
LZAKI   OPORL   OALAZ   IKIQI   QAHIW
CEAZN   AOFOX   QIZAG   ULOPE   XOOLI
DIDUG   MUVEQ   IIKFO   GUDIU   GIDUM
ECIDD   IANLO   MEQIZ   AUGHA   EJEJA
ZQICE   KINAG   UZANA   EVIQA   ZOLLA
KIIQK   IOLDI   CEAHO   RZAEC   HAANU
GIDNA   LO
```

Solution:—

Sighted Submarine Latitude Thirty Degrees Twenty One Minutes Longitude One Hundred Forty Two Degrees Fifteen Minutes At Nineteen Forty Fourteen March Period Possible This Is Base (Case?) Covering Great Circle Route To San Francisco

Keyword - Washington

Key Arrangement:-

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

```
MAPNC   HMDUS   YNLNN   PUSHC   YNFIN
YIFFI   PNFAH   CLHFT   PNCHO   CSUNP
NPFAY   OHLMT   TMIFE   PPNTM   MDYNO
PNYSU   USYNO   YHCPE   AFAML   ELNUT
PNYRH   LAFAF   LHFAR   YELDM   AMLEN
LMTLH   RYWSM   DWIPN   USLHY   NMANY
HLTMN   PSUCH   DMLEC   EPNCH   DMLED
MHLYN   AFLEC   HHCNY   NYFAE   LHCIW
SUELC   ODMLH   OCELU   TTMNP   ELAFM
ACOYN   POWIM   TNPMD   PNELM   TTMFA
PNCOP   NHCCH   LEUSL   NELUS   SULEA
FYRNP   OPPNF   ADMCH
```

Solution:—

Destroyers of offensive screen will fuel tomorrow Sunday beginning at daylight period Oilers take station as soon as practicable and complete all necessary arrangements. Plain component normal alphabet and mixed cipher component derived from a key using the letter frequency sequence from E to W resulting in the following values for letters not used in above text.

Plain J - Q - X - Z

Cipher UD - LR - WA - UI

ASSIGNMENT NO. 4 NOT RELEASABLE