Description of [illegible] Principles of an Investion (?) Machine for Locating Idiomorphs and Isomorphs in Cryptanalysis.

A. Preliminary Remarks

1. The experienced cryptanalyst very frequently has occasion to use the "probable word method", that is, he assumes the presence of a word in the text and tries to find it by one means or another. The word for which he is searching usually contains several repeated letters the identities and positions of which impart to the word a characteristic or peculiar "pattern" or "formula". Such words are termed idiomorphic. For example, the words BATTALION, DIVISION, and ARTILLERY are idiomorphic, and each has a distinctive formula of composition. The formula for the word BATTALION is representable by a sequence of arbitrarily-assigned digits standing for the different and the similar letters, as shown herewith:

B A T T A L I O N
1 2 3 3 2 4 5 6 7

In other words, the idiomorph is "coded" in the sense commonly spoken of in connection with the use of tabulating machinery operations.

2. Suppose the cryptanalyst suspects the presence of the word BATTALION in a message. Having established the formula for the word, he goes through the text, beginning with the first letter and sliding the formula against the text one step at a time, until he finds a place where the formula for a sequence of nine letters in the text is matched by or coincides with the formula 1 2 3 3 2 4 5 6 7 (= BATTALION). In this process the cryptanalyst must either indicate on paper, or see with his mind's eye, the textual formula that is at the moment in juxtaposition with the formula of the word for which he is searching.



3. Again, the cryptanalyst may not have a definite word in mind for which he is searching, but is simply desirous of locating two formulae which are identical and which are technically called isomorphs. In this case, he establishes a continuously changing series of formulae, beginning with the first letter of the text, and tests each of them against the remaining text of the same message or against the text of another message. For example, he may start with the first ten letters of the message and note the formula for these ten letters. He then applies this first formula to another part of the text, beginning at a given point and sliding the formula against this text one step at a time to see if he finds another sequence with an identical formula. If not, he starts with the second letter of the message, notes the formula for a sequence of ten letters, and repeats the search. When he encounters a formula which coincides with the one he has set up, he has found a case of isomorphism. The two isomorphs may or may not correspond to the same plain-text word. For example, the three words, WARRANT, LETTERS, and MISSION have identical formulae and are isomorphic.





14 April 1937

C 83.8

Filed: RAM - Comparators (1)

2

In searching for isomorphs, if the cipher text should contain all three of these words enciphered monoalphabetically, the cryptanalyst would not know whether the isomorphs he finds in the message represent three occurrences of the same word, or two occurrences of one word plus a different word with the same formula, or a single occurrence of three different words. These alternatives would be determined later, by further analysis. At the moment, the important fact is that he has located cases of isomorphism.

4. The present "hand" methods of locating idiomorphs and isomorphs are very tedious, slow, and subject to human error. Automatic machines for locating them would be quite useful, and it is the purpose of this invention to provide such machines and to describe the method of their use. Basically, the invention employs the comparison circuits used in the "coincidence counter" described in another paper, in combination with additional mechanism to be described.

5. The text in which idiomorphic or isomorphic sequences are to be located is prepared in the form of perforated Baudot tapes. Several copies of the text tapes are prepared, the number depending upon the length of the sequences for which a given machine is constructed. Machines for locating 6, 7, 8, ... letter sequences may be constructed. This description will deal with machines for locating sequences up to the length of ten letters. In general, the automatic locator goes through exactly the same steps as does the human operator at present, but at a much greater speed, and without overlooking possible places where sought-for sequences may exist.

## B. The idiomorph locator.

6. First to be described will be the idiomorph locator. Let us assume that we are searching for the word BATTALIONS, the presence of which is suspected in the following text:

A P D M N G F A A F V Q Z X R I Q V V W Z M A Z X Q O P A

7. The machine consists primarily of 10 Baudot tape transmitters, each being wired for comparison purposes, in a manner similar to that employed in the "coincidence counter". In this case, 10 tape copies of the message are prepared and these tapes are placed in the transmitters so that the successive starting points on the successive transmitters correspond to the successive letters A, P, D, M, ... of the message.

8. Closing a starting switch, the first letter of the message, A, in transmitter 1 is tested successively against the next 9 letters of the message, in transmitters 2, 3, 4, 5, 6, 7, 8, 9 and 10. This is done by means of a motor-driven rotary multiple switch, or other means, which successively connects the whole set of comparison contacts of transmitter 1 with those of transmitters 2 to 10, in turn.

The comparison circuits in this case merely serve to operate supervisory high-speed telephone relays to drive a series of 10 rotatable indicator wheels, the peripheries of which are divided up into 10 equal segments, numbered from 1 to 10. (This corresponds to the "coding" function referred to in Par. 1.) At the start all indicator wheels are at their reset or initial position, with segment 1 showing through a window or aligned at a bench mark. Indicator wheel number 1 is always set to position 1 and need not change; it may in fact be a "dummy" wheel. If no coincidence is encountered when transmitter 1 is compared with transmitter 2, indicator wheel 2 advances to segment 2; likewise, when transmitter 1 is compared with transmitters 3,4,5, ... and no coincidence is encountered in any case, indicator wheels 3,4, 5, ... advance to segment 2. But if a coincidence is encountered at any comparison, the indicator wheel associated with the transmitter being compared at that moment is not permitted to advance any further and remains in this position. The arrangement for this stepping may be by means of a continuously acting ratchet and pawl for each wheel and only when there is coincidence is the pawl withdrawn and locked into inactive position. For example, take the first 10 letters of the illustrative message:

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| A | P | D | M | M | G | F | A | A | F |

After transmitter 1 (set to A) has been tested against the rest of the transmitters, the indicator wheels are at the following positions:

| Indicator wheel | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Text | A | P | D | M | M | G | F | A | A | F |
| Setting of wheel | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 |

This corresponds with the repeated letter A in positions 1, 8, and 9 in the sequence A P D M M G F A A F. Indicator wheels 8 and 9 will now stay at position 1.

Transmitter 2 is now tested against the remaining transmitters, in turn. Since the letter P does not reappear in the sequence, the indicator wheels 3, 4, 5, 6, 7 and 10 advance one step. The formula is then:

| Indicator wheel | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Text | A | P | D | M | M | G | F | A | A | F |
| Setting of wheel | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 1 | 1 | 3 |

Transmitter 3 is now tested against the remaining transmitters, in turn. Since the letter D does not recur, the indicator wheels 4, 5, 6, 7, and 10 advance one step. The successive formulae set up on the indicator wheels as the successive transmitters are tested in turn against the remaining ones are as follows:

| Indicator wheel | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Text | A | P | D | M | M | G | F | A | A | F |
| (A) T1 against remaining T's: | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 |
| (P) T2 against remaining T's: | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 1 | 1 | 3 |
| (D) T3 against remaining T's: | 1 | 2 | 3 | 4 | 4 | 4 | 4 | 1 | 1 | 4 |
| (M) T4 against remaining T's: | 1 | 2 | 3 | 4 | 4 | 5 | 5 | 1 | 1 | 5 |
| (G) T6 against remaining T's: | 1 | 2 | 3 | 4 | 4 | 5 | 6 | 1 | 1 | 6 |
| (F) T7 against remaining T's: | 1 | 2 | 3 | 4 | 4 | 5 | 6 | 1 | 1 | 6 |

The formula for the word BATTALIONS is

| B | A | T | T | A | L | I | O | N | S |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 3 | 2 | 4 | 5 | 6 | 7 | 8 |

while that resulting from the test is 1 2 3 4 4 5 6 1 1 6. Therefore, the message does not start with this word.

9. The tapes in all transmitters are advanced one step and the same procedure is followed as before, now testing the 2d, 3d, ... 11th letters of the message. When the test begins with the 6th letter of the message, the following sequence of movements of indicator wheels occurs:

| Indicator wheel | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Text | G | F | A | A | F | V | Q | Z | X | R |
| (G) T1 against remaining T's: | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| (F) T2 against remaining T's: | 1 | 2 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 3 |
| (A) T3 against remaining T's: | 1 | 2 | 3 | 3 | 2 | 4 | 4 | 4 | 4 | 4 |
| (V) T6 against remaining T's: | 1 | 2 | 3 | 3 | 2 | 4 | 5 | 5 | 5 | 5 |
| (Q) T7 against remaining T's: | 1 | 2 | 3 | 3 | 2 | 4 | 5 | 6 | 6 | 6 |
| (Z) T8 against remaining T's: | 1 | 2 | 3 | 3 | 2 | 4 | 5 | 6 | 7 | 7 |
| (X) T9 against remaining T's: | 1 | 2 | 3 | 3 | 2 | 4 | 5 | 6 | 7 | 8 |

The final formula corresponds with that of BATTALIONS and hence the word BATTALIONS may exist at this point in the message.

10. It will be noted that the operator must stop after each test to compare the final formula with that of the word being sought. But by going one step further, the apparatus may be constructed so as to eliminate this source of delay.

11. Let the periphery of each indicator wheel be provided with projecting pins permuted in accordance with ten different combinations of the Baudot code and let these pins operate the series of contacts of a set of comparison circuits (each as in Converter Type M-134-T1) for testing the final formula against the formula for the word sought, as set up on a test frame, a circuit is completed which coincides with _____
Then, whenever the final formula of a text _____
lighted; otherwise the tape-stepping magnets of the tape transmitters are actuated and all the tapes stepped forward simultaneously whenever the final formula does not coincide with that set up on the test frame. Thus, once started on a message, the testing would continue without a stop until either a coincidence between a final formula and a test formula is encountered or else the end of the message is reached.

C. The isomorph locator.

12. In operating the idiomorph locator it is necessary to set up a formula on a test frame, against which successive final formulae are tested in turn. If the word sought for is not in the message, a new test formula has to be set up on the test frame and this would be done by hand. But in searching for isomorphs we do not have any specific formula in mind; we take any sequence of letters in the text and compare it with similar-length sequences in the same text to see if they are isomorphic.

13. Suppose that there are two sets of tape transmitters, each operating a set of indicator wheels and let one set, hereinafter called the test set of transmitters, be used merely to establish test formulae on the test frame, these formulae corresponding to and being set up by actual sequences in the message itself. Let the other set of transmitters, hereinafter called the message set of transmitters, be used as the source of successive final formulae to be tested against the formula set up on the test set of transmitters. Let the tapes on the message set be circular, by joining first and last characters. Let the arrangement be such that when a test formula is set up by the test set of transmitters it stays there until the entire message has been tested against it; if no coincidence has been encountered, then a new test formula is set up by the test set of transmitters by advancing the tapes in these transmitters, and the process repeated, the tapes in the message set of transmitters being back again at their initial positions. When a coincidence occurs between a formula on the test frame and a final formula set up by the message set of transmitters, a circuit is completed which lights a red lamp and stops all transmitters.

WILLIAM F. FRIEDMAN,
Principal Cryptanalyst,
Signal Intelligence Section,
War Plans and Training Division,
Office of the Chief Signal Officer.

Washington, D.C.
April 14, 1937.

The foregoing invention was described to us in February, 1937, by Mr. Friedman:

ROWLETT )
) Witnesses.
MILLER )

Cipher Machines

RAMS