From Bombe 'stops' to Enigma keys

A remarkably succinct description of the Bombe written many years ago, reads as follows:- '*The apparatus for breaking Enigma keys, by testing a crib and its implications for all possible settings and wheel orders*'. The accuracy of this description is not in question, but its brevity might lead to the false impression that the task of breaking the Enigma keys with the Bombe was a straightforward one carried out entirely by the machine; such a perception would be far from the truth. Before the Bombe could be used there were often tricky decisions to be made relating to the crib, and there were also certain procedures that had to be carried out by hand after the machine had completed its task, that could be time consuming and difficult.

The following notes attempt to describe some of the difficulties that could arise at different stages of the task and the methods that were employed to deal with them. No explanation of the basic principles of the Bombe is given, as several accounts are currently available elsewhere (e.g. BP Report No. 9 and Report No.16). It is assumed that the reader is familiar with the specification of the standard Enigma machine and of the German operational procedures used with it.

The Bombes were used to determine the following parts of an Enigma message key:-

(i) The rotor order.

(ii) The rotor core starting positions for the message.

(iii) The 'stecker' partner for a chosen letter.

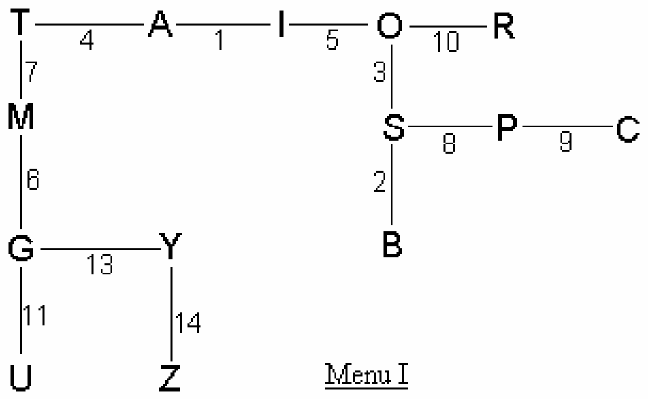
(It may be helpful to remind the reader that a complete set of steckers consisted of ten pairs of different letters from the alphabet, one letter in each pair being the stecker partner of the other. The Germans also imposed a rule whereby no pair ever consisted of two consecutive letters from the alphabet. The remaining six letters of the alphabet were said to be 'self-steckered' (in effect they were their own partners).

Before the Bombe could be used it was necessary to have obtained in some way a sequence of the plain-text characters corresponding to part of the cipher text. Such a sequence was known as a 'crib'. For example:-

Position:- 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Crib:- A B S T I M M S P R U Q Y Y …. ………..
Cipher text:- I S O A O G T P C O G N G Z N P M A J M B V Y

(ABSTIMMSPRUQYY…. was a frequently used wartime crib)

From this crib the following diagram known as a Bombe 'menu' can be drawn:-



Menu I

(The 12th link has been deliberately omitted in order to obtain a menu better suited for use with the following explanatory notes.)

An electric circuit representing the menu was ‘plugged up’ on the rear face of the Bombe, using the array of sockets provided, and sets of drums to emulate a possible Enigma rotor order was installed on the front face of the machine (when five different Enigma rotors were used there were 60 possible rotor orders). After the machine was started it carried out an electrical test at each of the 26×26×26 (=17,576) possible positions of the drums to identify those for which a possible stecker partner could be found for one letter that had been pre-selected from the menu. The Bombe was designed to stop automatically at any of the drum positions for which such a stecker partner was found. In particular a ‘stop’ would be expected to occur when the order and core positions of the drums on the Bombe were the same as the rotor order and the starting positions of the rotor cores originally used on the Enigma machine. In addition to this particular ‘stop’, additional *false* ‘stops’ would occur at other drum configurations, and these would sometimes cause difficulties, as there was no immediate way of distinguishing between the correct ‘stop’ referred to above, and all the false ones. (The false ‘stops’ were due to the inevitable consequences of chance, and not to technical faults on the Bombe).

Each ‘stop’ identified a particular drum order and the corresponding drum core positions together with a possible stecker partner for the letter that had been selected from the menu. By means of this information a set of possible stecker partners for all the other letters on the menu could be deduced as logical consequences.

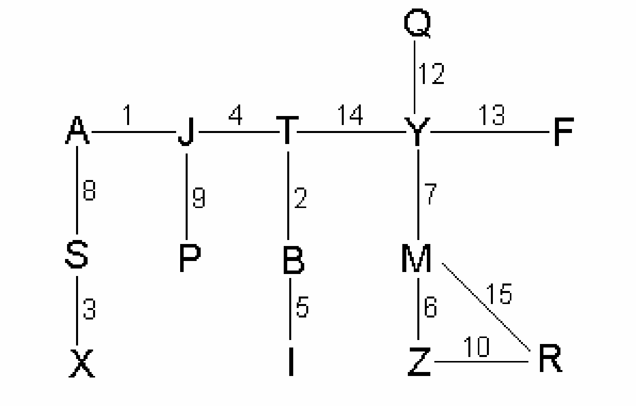
For any of the false ‘stops’ some of the deduced stecker partners would also certainly be false, but for the correct ‘stop’ all the deduced stecker partners would be correct. This distinction enabled the majority of the false stops to be identified and rejected by means of a routine procedure. However when the number of false stops was large then this routine procedure would become lengthy and time consuming.

The total number of false ‘stops’ generated from all the possible drum configurations depended entirely upon the structure of the menu, and one with many letters on it would be likely to give a smaller number of false stops than another with fewer letters. Menu I given above has 14 letters on it, and using it with a computer emulation of the Bombe the total number of stops obtained from sixty possible rotor orders was found to be 2,373. In this case the task of identifying the correct ‘stop’ would require the systematic elimination of well over two thousand false ones, and would be very demanding in terms of the human resources required, if it were to be attempted by hand.

Increasing the number of letters on the menu will reduce the number of false stops, but when a longer crib was available to make this possible, another difficulty was more likely to occur. In the Enigma machine the ‘turn-over’ positions of the middle and left-hand rotors are determined by the ring-settings used. These unknown settings could not be replicated on the Bombe, and consequently if any middle rotor ‘turn-overs’ had occurred during the original encipherment of the message, then the positions of the drums would not always match those of the rotors on the Enigma machine. Any menu would be invalidated if such a ‘turn-over’ had occurred at any position within its span, when all the ‘stops’ would be false.

A ‘turn-over’ of the middle rotor is bound to occur at some place within 26 consecutive positions of the right-hand rotor. Menu I spans 14 positions in the message, and hence the probability that it will be invalidated by a middle rotor ‘turn-over’ = 14/26. Using a longer menu with additional letters would increase further the probability of failure due to the occurrence of a middle rotor ‘turn-over’ within its span.

From the previous remarks it might appear that using a menu with fewer letters reduces the probability of such a failure but will incur the penalty of generating many more false 'stops'. Fortunately this is not necessarily true, as the number of false stops generated by a menu depends not only on the number of letters on it, but also on the arrangement of the links between them. This is demonstrated by the second menu:-



Menu II

Despite the fact that Menu II has the same number of letters as Menu I (both have14 letters), the total number of 'stops' generated from it is much smaller (101). This very significant reduction is the result of the presence in the menu of the 'closure' formed by the links between the letters M, Z, and R. Closures have the valuable effect of greatly reducing the number of random stops and were regarded as a highly desirable feature to have in any menu.

It must be remembered however that a menu could only be based upon a crib that was considered to have a good chance of being correct, and often this consideration resulted in one that did not have the desired characteristics. There was also the possible position of the middle rotor 'turn-over' to consider, which restricted the span of a plausible single menu. To quote from one authoritative source:- '*...it was largely a matter of chance which menu turned up trumps – so many things could go wrong , and no menu had a certainty of success. A 30 percent chance for a single menu or a 60 percent chance for a pair of menus linked to avoid the hazard of a middle rotor turn-over, was as good as you could normally hope for.*'

Alan Turing carried out a lengthy analysis to determine the number of stops per rotor order that could be expected from some different types of menu. The following table gives some of the results he obtained for menus consisting of a single web (like the two given examples).

ESTIMATED NUMBERS OF STOPS PER ROTOR ORDER

	Number of letters on the menu								
Closures	8	9	10	11	12	13	14	15	16
3	2.2	1.1	0.42	0.14	0.04	<0.01	<0.01	<0.01	<0.01
2	58	28	11	3.8	1.2	0.30	0.06	<0.01	<0.01
1	1500	720	280	100	31	7.7	1.6	0.28	0.04
0	40,000	19,000	7300	2,700	820	200	43	7.3	1.0

For Menu I the observed total number of stops was 2,373, equivalent to 2,373/60 ≈ 40 stops per rotor order. This menu contains 14 letters and had no closures so that the estimated value given by the table is 43.
For Menu II the observed total number of stops was 101, equivalent to 101/60 = 1.7 stops per rotor order. This menu contains 14 letters and 1 closure, so that the estimated value given by the table is 1.6. The two experimental results are sufficiently close to the estimated values to give reasonable confidence in the predictions given in the table.

Logistical difficulties would obviously arise if the number of false stops generated by a particular menu was large and so, except in special circumstances, it was standard practice not to use menus that were expected to give more than four stops per rotor order. The table shows that menus without any closures, normally had to have at least 16 letters on them, but for menus with one closure the minimum number of letters would be reduced to 14. A menu with two closures would require a minimum of 11 letters, and for three closures (a rare occurrence) a minimum of only 8 letters would be necessary.

As has been previously stated, most of the false 'stops' could be identified because for each of them at least some of the deductions for the possible stecker letter partners would be wrong and give rise to logical contradictions.
For example, suppose that two of the deductions arising from a 'stop' happened to be:- '*A is steckered to H*' and '*D is steckered to H*'. These logically contradict one another as no two letters on a menu can have the same stecker partner, and so that the 'stop' is false.
If all three letters were on the menu then the circuits in the Bombe would normally prevent the occurrence of any 'stop' caused by this type of contradiction, and in fact if one actually did occur it would be attributed to a technical fault on the machine. For this reason a 'stop' of this kind was said to be due to an '*illegal contradiction'.* Alternatively if two of the letters, say A and D, were on the menu but the third letter H was not, then the circuits in the Bombe could not prevent the 'stop' occurring, which was then said to be due to a '*legal contradiction*'. The identification of all the false 'stops' due to legal contradictions was carried out by hand with the aid of a small item of equipment known as a '*checking machine*'.

After the elimination of all the false 'stops' due to legal contradictions, there would remain a number of others, all with non-contradictory (i.e. confirmatory) sets of stecker partners for all the letters on the menu. With one exception, these stops would be false, for although each had survived the test procedure, the absence of any contradictory stecker partners was no guarantee that all the deductions were in fact correct, and for only one of these 'stops' would all the deductions be correct..
All of the 'stops' of this type were known as 'partial keys' (or 'stories') because if any one of them was set up on an Enigma machine it would decipher the letters in the message that matched those in the crib (this was known as 'deciphering the crib').

One of these partial keys or stories would also decipher the entire cipher message, but only after the complete set of steckers for it had been found. The task of identifying this particular story (i.e. the true one) and finding the additional correct steckers required to make up the complete the set for it, was often quite difficult and involved a degree of trial and error.

As an illustrative example, it will be recalled that Menu I gave a total of 2,373 stops, but after eliminating all those with legal contradictions between their steckers, sixteen partial keys or stories remained. Some of these stories were subsequently rejected because their deduced sets of steckers failed to conform to the procedural rules imposed by the Germans. In this way the number of partial keys was reduced to six (including the one as yet unidentified ‘true story’ that would lead to the Enigma key).

On some occasions during the war BP was obliged to attempt to find Enigma keys from cribs containing fewer letters than has so far been suggested. The table given previously shows that the numbers of false ‘stops’ will rise sharply when the number of letters on the menu is reduced to eleven or less. Under these circumstances the numbers of partial keys or ‘stories’ would also increase proportionately, and the task of identifying the true one would become very lengthy. A small number of more advanced version of the Bombe known as ‘Jumbo’ were constructed and these were capable of detecting all the ‘legal contradictions’, so that the only stops they made were for the partial keys (this explains how in the above example most of the false stops arising from Menu I were eliminated). Nevertheless with the weaker menus, even when a ‘Jumbo’ Bombe was employed, the resulting numbers of partial keys were too large to be conveniently dealt with in a realistic period of time by the human resources available.

Two papers in the BP archives that show that a mathematical investigation was carried out to find a very quick method for estimating the chance of a story being correct based only on the characteristics of the associated incomplete set of steckers (details are given in the appendix).

When the number of stories was small, the optimal order in which they were chosen for further detailed investigation could be decided from these estimates, so that the story with the greatest chance of being correct would be investigated first.

In cases when the number of stories was much larger, many of them could be rejected without the necessity of having to carry out the individual detailed investigations, because their estimated chances of being correct were found to be small.

In using such a procedure there was always the risk of rejecting the true story, but this risk had to be balanced against the impossibility of carrying out a detailed investigation of all of them. From the limited information currently available it is not clear to what extent this procedure was applied operationally, but in one original wartime paper there is a table of test values for the ‘goodness’ of any Bombe story, based entirely on the characteristics of its steckers, and a reference to an electrical device known as an ‘Analyser’ which had a display panel designed to show immediately the numerical test value given by the set of steckers given at each ‘stop’ by the more advanced version of the Bombe.

It would have been necessary to estimate the risk that was being taken by only carrying out a more detailed investigation of the ‘stops’ that gave test values above a chosen number (i.e. the risk that the correct ‘stop’ might be discarded). In common with many other statistical procedures developed at BP, these test values were based upon a logarithmic scale.

As an example, the six partial keys or stories derived from Menu I gave the results shown in the following table:-

Stop	No. self-steckers (s)	No. stecker pairs (r)	Test value (proportional to the odds of the stop being correct)
(a)	2	8	10
(b)	2	9	5
(c)	4	9	3
(d)	4	9	3
(e)	3	9	4
(f)	5	8	5

(These show that of the six stops, (a) is most likely to be the correct one.)
It so happens that this is true, but it must not be assumed that in all situations the stop with the greatest test value will invariably be the correct one. (Backing the favourite in a horse race is no guarantee that you have picked the winner.)

A brief and simplified description of the tricky task of identifying the true story, and augmenting the incomplete set of steckers to obtain the Enigma key, is given below.
It should be remembered that each of the stories selected for a more detailed investigation consisted of three parts:- (i) a possible rotor order, (ii) the initial positions of the three rotor cores, and (iii) a confirmatory but incomplete set of steckers.

Assumptions were made for the remaining unknown steckers, and the resulting conjecture of the Enigma key was set up on a British emulation of a Enigma machine to see if it would decipher parts of the complete message. The resulting sequence of letters would be examined to see (quote):- '*if there were signs of German clear text*' (i.e. additional small groups of letters outside the range of the crib that had appeared to have the characteristics of the German language. If there were such indications then appropriate changes would be made to the assumed steckers and the trial decipherment process repeated, until either the complete key was discovered or that particular story was rejected for an alternative one. This work was known as '*running a tape*', and clearly required considerable skill and patience. (There was a further complication relating to the determination of the original ring-settings, which has not been considered here.)

Frank Carter

====================

Mathematical Appendix

The likelihood that a particular 'stop' obtained from a Bombe run will provide the correct 'story'.

This depends on a number of factors each with its own level of certainty of being correct:-
(i) The crib used for the Bombe runs
(ii) The rotor order being tested.
(iii) The assumption made about the location of the middle rotor 'turn-over'
(iv) The characteristics of the steckers given by the stop.

The levels of certainty associated with the first three of these factors are unknown, but for a particular menu, will be the same for all the stops obtained. Consequently the likelihood of a particular stop being correct will be **proportional to** the odds that the corresponding incomplete set of steckers are part of the complete set of correct steckers.
Let the total number of Enigma plug-board combinations with 10 stecker pairs = **N**

$$(\mathbf{N = 1.507 \times 10^{14}})$$

Suppose that a given stop leads to **s** self steckers and **r** stecker pairs, and that some additional steckers are appended to them in order to make up a complete set, containing 10 stecker pairs, and six self steckers.

If **M(s, r)** represent the number of distinct ways in which this can be done, then the probability that the original **s** self steckers and **r** stecker pairs are correct = $\mathbf{M(s, r)/N}$

(This follows from the fact that one of the **N** possible complete sets must be correct)

The total number of rotor starting positions obtained from three chosen rotors = $\mathbf{26 \times 26 \times 26 = 26^3}$.
Hence the probability that the rotor core starting positions at the given stop are correct = $\mathbf{1/(26^3)}$
Let **S** represent the event that the given stop is correct, so that $\mathbf{Pr[S] = M(s, r)/N.26^3}$
(i.e. this is the probability that the given stop will provide the correct rotor starting positions and will also lead to **s** correct single steckers and **r** correct stecker pairs)
A **story** will decipher all the letters in the cipher message that are matched with the crib.
The probability that the correct stop will decipher all the letters on the crib = 1 (i.e. a certainty)
Let D represent the event that the given stop will decipher the letters on the crib.
then the probability that the given stop is correct and will decipher the letters on the crib:

$$\mathbf{Pr[S \text{ and } D] = \{M(s, r)/N.26^3\} \times 1}$$

The probability that the given stop is wrong: $\mathbf{Pr[S'] = 1 - \{M(s, r)/N.26^3\}}$

Suppose that the crib has **L** letters, so that the menu has **L** links, then the probability that a 'wrong' stop will decipher the letters in the crib = $\mathbf{(1/25)^L}$

(This is a consequence of the fact that no letter can be enciphered as itself)

Hence the probability that the given stop is wrong and will decipher the letters in the crib:

$$\mathbf{Pr[S' \text{ and } D] = \{1 - M(s, r)/N.26^3\} \times (1/25)^L}$$

$$= \mathbf{1 \times (1/25)^L} \text{ (approximately, since } \mathbf{M(s, r)/N.26^3} \text{ is small)}$$

O(S | D) represent the odds that the stop is correct given that it has deciphered the letters in the crib.

Then $\mathbf{O(S \mid D) = Pr[S \mid D]/Pr[S \mid D]'}$ (by definition)

$$= \frac{\mathbf{Pr[S \text{ and } D]/Pr[D]}}{\mathbf{1 - \{Pr[S \text{ and } D]/Pr[D]\}}}$$

$$\text{hence } \mathbf{O[S \mid D]} = \frac{\mathbf{Pr[S \text{ and } D]}}{\mathbf{Pr[D] - Pr[S \text{ and } D]}} = \frac{\mathbf{Pr[S \text{ and } D]}}{\mathbf{Pr[S' \text{and } D]}}$$

Substituting the expressions previously derived for these two probabilities:-

$$\mathbf{O[S| D] = [M(s, r)/N.26^3] \times 25^L}$$

Hence the odds on a particular stop being correct are proportional to the value of the expression:-

$$\mathbf{X = \frac{M(s, r) \times 25^L}{N.26^3}} . \qquad \text{It can be shown that } \mathbf{M(s, r) = \frac{\{26 - (s + 2r)\}!}{(6 - s)!\, 2^{10-r}\, (10 - r)!}}$$

(Where $\mathbf{s \le 6}$ and $\mathbf{r \le 10}$)

A logarithmic scale (to base 10) was originally adopted, so that the final value used and denoted by the letter **C** was: $\mathbf{C = 5 \times \log (X)}$
Two examples to show how the results in the table were obtained:-
Menu I shown previously has 13 links (i.e. L = 13).

Stop 'a' has 2 self-steckers and 8 stecker pairs (i.e. s = 2, r = 8, L = 13)
For this stop M(2, 8) = $8!/(4! \times 2^2 \times 2!)$ = 210
Hence the odds on stop A being correct are proportional to:-

$(210 \times 25^{13})/(1.507 \times 10^{14} \times 26^3) \approx 118$

Hence C = 5 × log (118) = 10.3 (≈ 10)

Stops 'c' and 'd' have with 4 self-steckers and 9 stecker pairs (i.e. s = 4, r = 9, and L = 13)
For this stop M(4, 9) = $4!/(2! \times 2^1 \times 1!)$ = 6
Hence the odds on stop 'c' being correct are proportional to:-

$(6 \times 25^{13})/(1.507 \times 10^{14} \times 26^3) \approx 3.37$

Hence C = 5 × log (3.37) = 2.6 (≈ 3)
