CS341 ps02

July 12, 2022

Question 3

Prove that the signature of an Enigma configuration is independent of the plugboard. (I.e., argue that the signature of a particular rotor sequence/setting is unchanged when you alter plugboard pairings.)

Definitions
Cycle: A cycle is a set of vertices connected in a closed chain. In our case, these cycles are components of a 26-node directed graph, where each node represents a different letter of the English alphabet. An edge of this graph between nodes u and v denotes "if l is the letter that encrypts as u in the base configuration, then l encrypts as v in the base + 3 configuration."

Signature: Let's call the signature of a particular base setting of the Enigma machine the profile of lengths of these cycles in descending order of length.

Constant Enigma Configuration: The slow, medium, and fast rotors are already chosen and remain constant along with their initial positions.

Proof by Induction
We will prove that the signature of an Enigma configuration is independent of the plugboard through a proof by induction.

Base Case:
We will first prove that the signature of a constant Enigma configuration with no plugboard pairings (n = 0) remains the same (Note: we omit the use of ring settings for this proof).

At a particular index of a constant Enigma configuration with no plugboard pairings, a certain input letter will always map to a certain output letter.
Otherwise, the Enigma machine would be broken and might not decode properly.

Each edge in the graph of cycles used to compute the signature will remain constant

This is because an arbitrary letter, "A", must always map to a particular letter in the base index and a particular letter in the base+3 index.

The graph used to compute the signature will only contain cycles, and these cycles remain constant

There are a finite number of letters in the alphabet, 26; there exists an edge for each letter in the alphabet in the base index to a letter in the base+3 index; no two letters map to the same letter in the same index; and each edge between letters will remain the same. Thus, only non-overlapping cycles can exist under these conditions.

Therefore, the signature of a constant Enigma configuration with no plugboard pairings (n = 0) remains the same

Because these cycles are constant and non-overlapping, their lengths remain the same. Since there exists only one signature for a set of cycles of the same lengths, the signature remains constant.

Inductive Case

We want to prove that if the signature of a constant Enigma configuration with n plugboard pairings is constant, then the signature of that same Enigma configuration with n + 1 plugboard pairings is equal to the signature of that same configuration with n plugboard pairings.

Assumption:

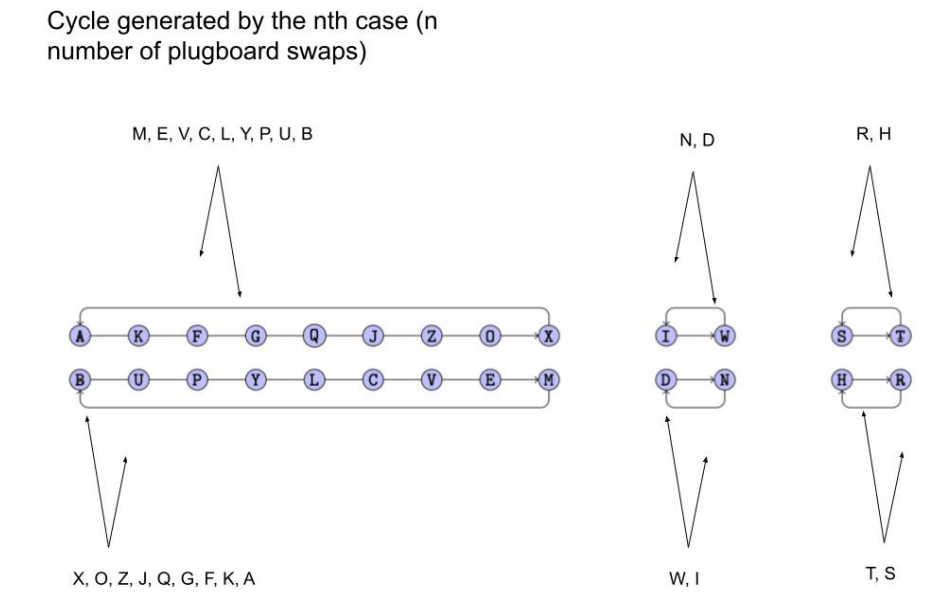
We assume that a constant Enigma configuration with any n plugboard pairings has a constant signature. Note the practical maximum of n is 13, due to the constraints of the alphabet.

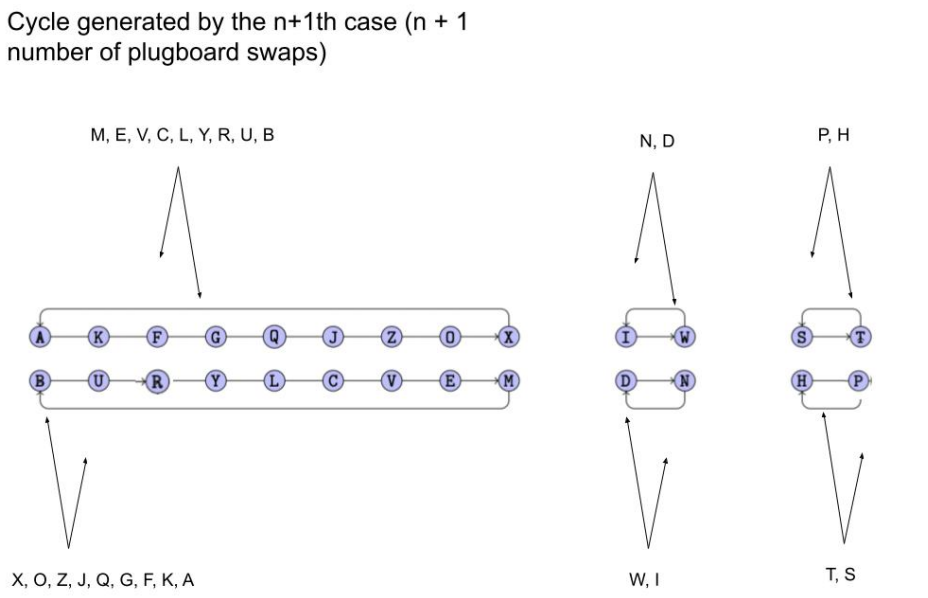
Only two characters are involved in a certain plugboard pairing.

This is based on the wiring of the machine.

Adding a plugboard pairing between two arbitrary letters "A" and "B" has the effect of swapping the letters "A" and "B" within the sets of letters that map to individual cycles and swapping the letters "A" and "B" within the nodes of the cycles themselves.

This is because a swap between these two letters causes encrypting the letter "B" instead of "A" whenever "A" is pressed, and as well as whenever "B" is supposed to be outputted letter "A" is outputted instead. The equal but opposite relationship holds when "B" is pressed and when "A" is supposed to be outputted.





Figures showing the cycles generated by n and n+1 number of plugboard swaps, where the n+1th plugboard swap is between the letters P and R (bottom left and bottom right cycles)

Adding any individual plugboard pairing between any two arbitrary letters "A" and "B" does not change the length of the cycles nor the number of letters that map to nodes in a particular cycle.

Thus, the signature of that same constant Enigma configuration with n + 1 plugboard pairings is equal to the signature of that same configuration with n plugboard pairings.

Combined, the fact that the signature of a constant Enigma configuration remains the same when there are no plugboard pairings and we can always add any plugboard pairing and get the same signature means that **the signature is independent of the plugboard.**

Question 5

Ignoring ring settings and plugboard pairings, there are still over one million (1,054,560 exactly) different configurations with just the rotors and starting messages. However, even getting the most common signature, being (13, 13), will reduce the amount of possibilities to around $\frac{1}{4}$ of the original settings (to exactly 262,462). The mean number of configurations per signature is approximately 11,717, and therefore reduces the possibilities for a given day by over 98% on average. Therefore, this approach significantly helps identify the correct Enigma configuration.

Other measures of the average number of configurations per signature include the mode and median. The mode, being the most common number of configurations per signature, is 1 and 2. The middle value, the median, is 1198. Therefore, the median is less than the mean, signifying that the distribution is right-skewed; there are many signatures with a small number of configurations, and a few signatures with a large number of configurations.

Although this approach narrows down the space of possible Enigma configurations for a given day, the plugboard pairings specifically still complicate the linguist's work. Since the signature is independent of the plugboard pairings, this approach still leaves a lot of work for the linguists/analysts to reduce the data set even further because every possible plugboard pairing is still possible with every configuration analyzed with this approach.