Comput 331 Assignment 1 Part 4

4. a: How many distinct keys, keys which each produce different ciphertexts for the same message, do each of the four ciphers have?

- The original will have 25 distinct keys, one for each possible shift based on the alphabet.
- The Part 1 cipher will have 51 distinct keys since it includes all upper-case and lower-case alphabet, unlike the original version which converts all messages to upper-case.
- The Part 2 cipher will have 52 distinct keys since after the first one was shifted, the following character will encrypt based on the previous letter.
- The Part 3 cipher will have 52^n distinct keys, where n is the length of the message, in other words, it can reach infinite since the length of the key can be infinite.

4. b: Is the problem 2 cipher stronger than the original Caesar cipher?

- Yes, the problem 2 cipher is stronger than the original Caesar cipher, since the original version can be brutally forced easily with 25 iterations of different shift amounts. However, the second problem needs at most 52 iterations to be cracked, since the first letter of the message can be shifted with any of the 52 alphabets.

4. c: Can you think of any weaknesses your Caesar Cipher modifications have that do not pertain to the number of keys and that may allow an attacker to easily guess your key? Discuss how this weakness could be addressed. (Hint: Think about lowercase and uppercase letters).

- Lowercase and uppercase letters in Caesar Cipher can be a big problem since hackers can use the frequency of letters in the encrypted message to find commonly used letters, then based on its uppercase or lowercase check if the Cipher uses fixed shift amount. For example, if they figure that ‘e’ is encrypted to ‘g’, and ‘E’ is also encrypted to “G”, then they can find out the fixed shift amount to be fixed, and use that to crack the code easily. Basically, in our modified Caesar's Cipher algorithm, the order of the letters is an enormous problem. A is always between a, and E is always between e, thus we only need to randomize the letters with their Capital and their position in the letters makes it more difficult to decrypt.
- Also, since the frequency of letters in encrypted text can be very useful when comes to decrypting due to the nature of English, we can use multiple letters to represent the same letter in plain text. For example, if the letter ‘a’ occupies 10% of all the letters in literary English, we can come up with 10 letters that can represent an encrypted text. There ‘a’ can be represented by one of the 10 letters, which will bring each letter occupation in the text to 1%. We can create a new letter table based on it to break the characteristic of literary English.