**Homework #1**

Due date: **27/10/2023**

Notes:

* Your work (code + written answers) must be submitted through SUcourse+.
* Winzip your programs and add a readme.txt document (**if necessary**) to explain the programs and how to use them.
* Name your winzip file as **“cs411\_507\_hw01\_yourname.zip”**

1. (**15 pts**) Consider the shift cipher. Show that the ciphertext “NLPDLC” can be decrypted into a meaningful word. Find out this word and the corresponding encryption key.

To solve this problem, we employed the shift cipher decryption technique, also known as the Caesar cipher. This method entails shifting each letter of the ciphertext by a fixed number of positions. We began by attempting all possible shift values, ranging from 1 to 25, to ascertain if any of the decrypted versions yielded a meaningful word.

Upon analysis, it became evident that the letter 'L' occurred most frequently in the ciphertext 'NLPDLC'. This led us to deduce that the 'L's in the ciphertext likely corresponded to 'E's in the plaintext, given that 'E' is the most commonly used letter in the English alphabet. Employing the formula 'Y = a + k mod 26', where 'a' represents the index of 'E' and 'Y' represents the index of 'L', we derived a shift value of 7.

However, when applying this shift, we obtained the decrypted text 'GEIWEV', which did not yield a meaningful word. Undeterred, we then considered the second most frequent letter, 'A'. Applying the same formula with 'a' as the index of 'A' and 'Y' as the index of 'L', we obtained a shift value of 11.

With a shift value of 11, we successfully decrypted the ciphertext to reveal the meaningful word 'CAESAR'. Hence, our key is 11, and the plaintext is 'CAESAR'."

1. (**20 pts**) Consider the ciphertext generated by Affine Cipher over Z26. As a hint, you are told that the most frequent letter in the plaintext is ‘T’. Find the plaintext, the encryption and decryption keys. Show your work.

"J gjg mxa czjq ayr arpa. J ulpa cxlmg ayerr ylmgerg rqrwrm hzdp ax gx ja hexmn."

To approach this problem, we began by examining the ciphertext and conducted a frequency analysis to determine the most commonly occurring letter, which turned out to be 'A'. This information served as a starting point. Given that 'T' is the most frequently used letter in English, we made an educated assumption that 'T' in the plaintext corresponds to 'A' in the ciphertext.

Next, we applied the principles of the Affine Cipher, using the equation 'y = ax + b mod 26' where 'a' and 'b' represent the encryption key, and 'x' and 'y' are within the range of 0-26. By substituting the indices of 'A' and 'T' into this equation, we derived 'a = 11' and 'b = 25' as suitable values.

With these encryption keys in hand, we employed a provided code to decrypt the message. This code utilized the Affine Cipher decryption process, which involves finding the modular inverse of 'a' and executing the necessary calculations. The result confirmed our expectations, revealing the original message: 'I DID NOT FAIL THE TEST. I JUST FOUND THREE HUNDRED ELEVEN WAYS TO DO IT WRONG'.

In summary, the encryption key is 'a = 11' and 'b = 25', while the corresponding decryption key is 'a\_inv = 19' and 'b = 25'.

1. **(15 pts)** Assume that you design a new affine cipher where you encrypt two letters at a time, where your alphabet is

{'A':0, 'B':1, 'C':2, 'D':3, 'E':4, 'F':5, 'G':6, 'H':7, 'I':8, 'J':9, 'K':10, 'L':11, 'M':12, 'N':13, 'O':14, 'P':15, 'Q':16, 'R':17, 'S':18, 'T':19, 'U':20, 'V':21, 'W':22, 'X':23, 'Y':24, 'Z':25, '.':26, ' ':27, ',':28, '!':29}.

In other words, you group your plaintext message in bigrams (i.e., two-character words) and encrypt each bigram of the plaintext separately using this affine cipher. For example, if the first two letters of a plaintext is “TH” then it will be encoded as follows

TH ⇒ 19 × 30 + 7 = 577.

If the number of letters in the plaintext is not a multiple of two, you pad it with the letter “X” at the end. Determine the modulus and the size of the key space.

The modulus for this given alphabet is 900, signifying that there are 30 possible options for both the first and last letters in the bigrams. To determine the size of the key space, we need to find the number of 'a' values where the greatest common divisor (gcd) with 900 is 1. This process led us to a total of 240 'a' values, as determined using our gcd() function in the provided code (located in the zip file). Multiplying this by the modulus of 900 yields a key space size of 216,000.

1. (**15 pts**) Is the affine cipher defined in question (3) secure against the letter frequency analysis?

The security of this affine cipher against letter frequency analysis lies in the dynamic nature of its key generation. The arrangement of letters within a bigram dictates the key used for encryption. For instance, if we consider the ciphertext 'HATAAH', 'A' emerges as the most frequently occurring letter. It might be tempting to assume that 'A' in the ciphertext corresponds to 'E' in the plaintext through letter frequency analysis (LFA). However, it's crucial to note that each 'A' possesses a distinct encryption key, contingent on the order of letters within the bigram, or even the presence of the other letter. As a result, applying LFA in this context would not yield a reliable outcome.

1. **(20 pts**) Consider the following ciphertext that is encrypted with the affine cipher defined in question (3):

"ZHOFC.BNZCLRZ WNJ.XGI.WMBDV.MEJ!GGYKGDZ ERGMWNJ.KDGD RSW"

Find the key and decrypt the ciphertext.

(**Hint 1:** The plaintext is a sentence that ends with a dot.)

(**Hint 2:** The length of the plaintext (plen) is not a multiple of 2; plen = 2k+1 for an integer k.)

To decrypt the given ciphertext, we first need to find the modular inverse of the coefficient 'a', which is the number used in the encryption process. In this case, 'a' is 19. The modular inverse of 'a' (notated as (a^{-1})) is an integer that, when multiplied by 'a', gives a result of 1 when taken modulo 30. In this case, the modular inverse of 19 is also 19, as (19 dot 19 mod 30 = 1).

Next, we find the modular inverse of the additive constant 'b', which is 7. The modular inverse of 'b' (notated as (b^{-1})) is the number that, when multiplied by 'b', gives a result of 1 when taken modulo 30. In this case, the modular inverse of 7 is 23, as (7 dot 23 mod 30 = 1).

With the modular inverses calculated, we proceed to decrypt the ciphertext. We do this by breaking the text into pairs of two characters, known as bigrams. Each bigram is decrypted separately using the formula (x = a^{-1}(y - b) mod 30), where 'y' is the numerical representation of the character.

For example, the bigram "ZH" is decrypted as (8 \* 23 + 23 = 207), which corresponds to the character "H". This process is repeated for each bigram in the ciphertext, resulting in a partially decrypted message.

Since the length of the plaintext is not a multiple of 2, one extra character is added to the end of the message to form a bigram. In this case, the letter "X" is used for padding. This extra character is identified and replaced with the correct character.

After decrypting all the bigrams and correcting for padding, we arrive at the final decrypted message: "SING, GODDESS, OF THE ANGER OF ACHILLES, SON OF PELEUS." This is the complete and accurate decryption of the original ciphertext with the key a=91, b=389.

1. **(15 pts)** If we select a different shift amount for every letter in the plaintext uniformly randomly, the shift cipher becomes a one-time-pad with perfect security. Suppose pα is the probability of the plaintext letter α from the Turkish alphabet, where α ∈ {A, B, C, Ç, …, Z}. Suppose also that pβ is the probability of the ciphertext letter β, where β ∈ {A, B, C, Ç, …, Z}. Demonstrate that pβ= 1/29 for every β ∈ {A, B, C, Ç, …, Z} independent of the values of pα.  
     
   In this scenario, we're considering a situation where each letter in the plaintext is shifted by a randomly chosen amount, effectively transforming the shift cipher into a one-time pad with perfect security. We aim to demonstrate that, regardless of the probabilities assigned to individual plaintext letters (p\_α), the probability of any given ciphertext letter (p\_β) is consistently 1/29, where β can be any letter from the set {A, B, …, Z, ., , ,, !}.

The rationale behind this lies in the uniform randomness of the shift amounts. This ensures that predicting the next letter in the ciphertext is entirely unbiased. For example, if we shift 'A' by 7, 'B' by 2, and 'C' by 27, the distribution of shift amounts provides no hint about what the shift amount for 'D' might be. This means that up until 'D', the distribution of shift amounts for each letter remains distinct and unbiased, resulting in a state of uniform randomness. Consequently, the probability of any subsequent ciphertext letter being a specific 'β' is consistently 1/29.

1. **BONUS (20 pts)** The following was encrypted using the Vigènere cipher:

“FNZ FFZZMLQQZVO GAXXH PZ UPU QXGIHU UY NWJXR AHBDLPOMK YOUPZM, VOZAYCD. J TGQH B XUIJJZM ARS XOAH, BZJ D JP AT GLWUTB LO EVDWF AL GRHUI. OKPGMC L NME IRU NKGLFHK DQ UTK JUEQX JI UTK PQJHKMVF, KKO L MABZ WIQ YOLDWE GLUFRZ OFMBZV BE ZCHZ AVZQ JZ YKUJZM. D OPHK OKF NRPH TWE, D OPHK NRNQ VZRQXK, RKPY UIH MABZV ZAA FQPI YJPFFOHHT IOOKPGZ FQPIOIJ XTE. D OPHK NRNQ MMHBF JZHEE JJQF NE HHO, FNJXHT O’QH MATB FFMYZG QQXCDQE ZJ KBHK ADJFN DQ UTKH, BFF LMRN ARY KBNOO ROQ’Y CHBDZ KUJLKN WIQS. CHSQ ZCHZ TGQH CDUPJIF ZCH TAAK IPD EJX, FMZ DW, JF CDOM PU TRV SUJG. JF’Y ALSEZ-MDUQ YJXQ, FNZB LZUR KPI ZJ PBWK DW IQXZ. L XMTO WP FXVYFX OI HVDUKH, BXEJVIM, O NKBXR NHU ALA ISAS CHSQ. GIG ZQZ D NOAC OKBF O VP PZRT JPUTB WP M MMDWQEVUE, NAO LU’E G HRTF VMHDUUPV HDGQHZMXY, WIMZ’N ZIMZ DW JE. VMHDUUPV BDK OKF PKVG UTGO OJQ ZCHSQ, KQHSK YOROQ UQHS FNZP TBKVNT AL NXDT HPUOUTB OJRK DQ UTK KDTF, UA VVON KDTEOJQBFK ADJFN DQ UTKDU XAXF, WIQOM WSGZC, WIQOM VUDABJMQ GIG UTKDU TOOZQDQ, ZCDU U QIRX U YCDMX LVOM AT OKF SXJXOP GIG LUYN WIAYZ VUATZV BZJ RHFB UQHS FNZP; UTUPJI U’S XROHOIFFP OI PZ TKVUU FNVW JF’Y GROS HZHO ZUOKJZM WXU M MMDWQEVUE. MTY L TTGGO OAZ RHFB LMRN PKNSBUX, WXU EOHSMK HZFBGYZ L TTGGO CQ NVSQK OI PZ FKVUT, U YCDMX YOHFB ST VPGR DQ FYUOLPZ. O GRWQ ZCH TFOXNZ XKVYFE OI VQDOIJ, UTK WOVQ YFB - UTGO’V BXR DW JE. OO’V OAZ V PBFZZU PR OIWFXRZFU AX GRHUI, DW’T XUQLOS CDWI ATZ’V JZYDGF, IOOK PZK’N VUASVFI.”

Attack it and find the key length and the key. Note that only the letter characters are encrypted.