Quiz 2

Status	Done
Course	🖥️ 密碼工程
Due date	@March 9, 2023
Paper	
Type	

1. **Please determine the dimension of the rectangle for this encryption cipher.**

Sol. The dimension of the ciphertext is 9 rows x 7 columns.

7 rows x 9 columns

```
EALESVTRA
CEEROBIIA
DRDNDNQGE
TASEYLLAI
MUANTCAWH
EOMANREEO
COMSRLTBR
```

Number of vowel

```
4
6
1
4
3
6
1
```

Difference (- 9 x 0.4 = 3.6)

```
0.4
2.4
2.6
0.4
0.6
2.4
2.6
```

💡 The sum of differences with 7 x 9 rectangle = 11.4

9 rows x 7 columns

```
ERASBLE
CAMSNAB
DUMOLEA
TOEDCTA
MORYRRE
ELNTLII
CEENTGH
ADNRIAO
ESAVQWR
```

Number of vowels

```
3
2
4
3
2
3
2
4
2
```

Difference (- 7 x 0.4 = 2.8)

```
0.2
0.8
1.2
0.2
0.8
0.2
0.8
1.2
0.8
```

💡 The sum of differences with 9 x 7 rectangle = 6.2

By the above result, sum of differences of 9 x 7 rectangle is smaller than 7 x 9 rectangle, so the dimension of the rectangle is more likely to be 9 x 7.

```
ciphertext1 = [
"ERASBLE",
"CAMSNAB",
"DUMOLEA",
"TOEDCTA",
"MORYRRE",
"ELNTLII",
"CEENTGH",
"ADNRIAO",
"ESAVQWR",
]
ciphertext1_fake = [
"EALESVTRA",
"CEEROBIIA",
"DRDNDNQGE",
"TASEYLLAI",
"MUANTCAWH",
"EOMANREEO",
```

```
"COMSRLTBR",
]
def print_dimension(ciphertext):
    print(f"The dimension of the ciphertext is {len(ciphertext)} x {len(ciphertext[0])}")

def count_vowels(ciphertext):
    sum = 0
    for i in range(len(ciphertext)):
        vowels = 0
        for ch in ciphertext[i]:
            if ch == 'A' or ch =='E' or ch =='I' or ch =='O' or ch =='U':
                vowels += 1
        print("vowels for", ciphertext[i], "is", vowels, ", and differences is", abs(round(vowels - len(ciphertext[0]) * 0.4, 2)))
        sum += abs(round(vowels - len(ciphertext[0]) * 0.4, 2))
    print("The sum of differences for")
    pprint(ciphertext)
    print("is:", sum)
    print("---------------")
    return sum

print("Answer to the question 1:")
sum1 = count_vowels(ciphertext1)
sum2 = count_vowels(ciphertext1_fake)
print("Since", sum1, "<", sum2)
print_dimension(ciphertext1)
```

以上是我第一題的code，由於老師在上課有提示說這個密文不是7x9就是9x7, 因此我只有考慮7x9以及9x7的狀況。

我先將密文依照第二題的hint的方向塞進ciphertext1以及ciphertext1_fake裡面，
接著寫出兩個函式，分別是

1. print_dimension，用途是將這個密文的矩陣大小算出來。
2. count_vowels，用途是將每個row的vowels次數算出來以及將每個row的vowels differences算出來，並將sum of differences算出來。

最後我再人工判斷發現ciphertext1_fake比ciphertext1的sum of differences比較大，所以print出兩者的大小關係，並將dimension print出來。

```
Answer to the question 1:
vowels for ERASBLE is 3 , and differences is 0.2
vowels for CAMSNAB is 2 , and differences is 0.8
vowels for DUMOLEA is 4 , and differences is 1.2
vowels for TOEDCTA is 3 , and differences is 0.2
vowels for MORYRRE is 2 , and differences is 0.8
vowels for ELNTLII is 3 , and differences is 0.2
vowels for CEENTGH is 2 , and differences is 0.8
vowels for ADNRIAO is 4 , and differences is 1.2
vowels for ESAVQWR is 2 , and differences is 0.8
The sum of differences for
['ERASBLE',
 'CAMSNAB',
 'DUMOLEA',
 'TOEDCTA',
 'MORYRRE',
 'ELNTLII',
 'CEENTGH',
 'ADNRIAO',
 'ESAVQWR']
is: 6.2
---------------
vowels for EALESVTRA is 4 , and differences is 0.4
vowels for CEEROBIIA is 6 , and differences is 2.4
vowels for DRDNDNQGE is 1 , and differences is 2.6
vowels for TASEYLLAI is 4 , and differences is 0.4
vowels for MUANTCAWH is 3 , and differences is 0.6
vowels for EOMANREEO is 6 , and differences is 2.4
vowels for COMSRLTBR is 1 , and differences is 2.6
The sum of differences for
['EALESVTRA',
 'CEEROBIIA',
 'DRDNDNQGE',
 'TASEYLLAI',
 'MUANTCAWH',
 'EOMANREEO',
 'COMSRLTBR']
is: 11.4
---------------
Since 6.2 < 11.4
The dimension of the ciphertext is 9 x 7
```

以上是我的第一題輸出結果。

2. **Please Solve this following transposition cipher which involves a completely filled rectangles from the HINT below.**

Sol.

```
def column_swap(ciphertext, key):
    tmp = [list(row) for row in ciphertext]
    for i in range(len(ciphertext)):
        for j, k in enumerate(key):
           tmp[i][j] = ciphertext[i][k - 1]
    return tmp
str_array = column_swap(ciphertext1, [6, 3, 4, 1, 2, 5, 7])
ciphertext1 = ["".join(row) for row in str_array]
print("========================")
print("Answer to the question 2:")
pprint(ciphertext1)
```

以上是我的第二題的code。

我寫出了一個函式column_swap，目的是將原本的密文做column之間的對調以解出原本的明文，在嘗試幾次以後我發現透過6341257這把key就能將原本的明文變出來。

原本的明文是:
LASER BEAMS CAN BE MODULATED TO CARRY MORE INTELLIGENCE THAN RADIO WAVES

另外我認為最後的QR是拿來補足矩陣大小用的，沒有意義。

```
 'LASERBE'
 'AMSCANB'
 'EMODULA'
 'TEDTOCA'
 'RRYMORE'
 'INTELLI'
 'GENCETH'
 'ANRADIO'
 'WAVESQR'
```

以上是我的第二題輸出結果。

3. **Please count Index of Coincidence (IC) for each message. The IC of English is around 0.66**

Sol.

```
message1 = """CRYPTANALYSIS IN RECENT PUBLICATIONS ALSO CRYPTANALYSIS
REFERS IN THE ORIGINAL SENSE TO THE STUDY OF METHODS AND
TECHNIQUES TO OBTAIN INFORMATION FROM SEALED TEXTS THIS
INFORMATION CAN BE BOTH THE KEY USED AND THE ORIGINAL TEXT
NOWADAYS, THE TERM CRYPTANALYSIS MORE GENERALLY REFERS TO
THE ANALYSIS OF CRYPTOGRAPHIC METHODS NOT ONLY FOR CLOSURE
WITH THE AIM OF EITHER BREAKING THEM I E ABOLISHING THEIR
PROTECTIVE FUNCTION OR OR TO PROVE AND QUANTIFY THEIR
SECURITY CRYPTANALYSIS IS THUS THE COUNTERPART TO
CRYPTOGRAPHY BOTH ARE SUBFIELDS OF CRYPTOLOGY
"""

message2 = """DIE KRYPTOANALYSE IN NEUEREN PUBLIKATIONEN AUCH
KRYPTANALYSE BEZEICHNET IM URSPRUNGLICHEN SINNE DAS STUDIUM
VON METHODEN UND TECHNIKEN UM INFORMATIONEN AUS
VERSCHLUSSELTEN TEXTEN ZU GEWINNEN DIESE INFORMATIONEN
KONNEN SOWOHL DER VERWENDETE SCHLUSSEL ALS AUCH DER
ORIGINALTEXT SEIN HEUTZUTAGE BEZEICHNET DER BEGRIFF
KRYPTOANALYSE ALLGEMEINER DIE ANALYSE VON KRYPTOGRAPHISCHEN
VERFAHREN NICHT NUR ZUR VERSCHLUSSELUNG MIT DEM ZIEL DIESE
ENTWEDER ZU BRECHEN D H IHRE SCHUTZFUNKTION AUFZUHEBEN BZW
ZU UMGEHEN ODER IHRE SICHERHEIT NACHZUWEISEN UND ZU
QUANTIFIZIEREN KRYPTOANALYSE IST DAMIT DAS GEGENSTUCK ZUR
KRYPTOGRAPHIE BEIDE SIND TEILGEBIETE DER KRYPTOLOGIE
"""
```

```
message3 = """MVWZXYXEJIWGC ML BIAORR ZYZVMAKXGYRQ KPQY GPITRKRYVCQSW
POJCBW GX XFO SPSKGXEJ CILCI RY XFO WREHW YJ KOXFYHQ KRB
DIARRGAYCC XM YFRKML SRDYVKKXGYR DBSK CIYVIB DIVDW RRMQ
SRDYVKKXGYR AKR ZO FMDL RRI IOC SCIB KRB DLC YVGQMLKP ROBR
XSUKHYIW, RRI ROVK MVWZXYXEJIWGC QMBI EORCBEJVC POJCBW RY
XFO ELKPWCMQ YJ ABCNDSEBENRMA WIRRSBC RMD SLVC DYV AVSQEVC
GMRR XFO EGW SD OMRRIP LVCKOGXK RRIK S I YLSJSWFSRE DLCSV
NBSROGRSZC PYLMXGYR MB SP DS NBSTO ELN USKRRSJW DLCSV
QOGSBMRI GPITRKRYVCQSW GC XFEW RRI AYYLDIPZEPD XM
MVWZXMQVYZLW LSRR EPO WSLJGOPBC SD MVWZXMVSEI"""

message4 = """FUBSWDQDOBVLV LQ UHFHQW SXEOLFDWLRQV DOVR FUBSWDQDOBVLV
UHIHUV LQ WKH RULJLQDO VHQVH WR WKH VWXGB RI PHWKRGV DQG
WHFKQLTXHV WR REWDLQ LQIRUPDWLRQ IURP VHDOHG WHAWV WKLV
LQIRUPDWLRQ FDQ EH ERWK WKH NHB XVHG DQG WKH RULJLQDO WHAW
QRZDGDBV, WKH WHUP FUBSWDQDOBVLV PRUH JHQHUDOOB UHIHUV WR
WKH DQDOBVLV RI FUBSWRJUDSKLF PHWKRGV QRW RQOB IRU FORVXUH
ZLWK WKH DLP RI HLWKHU EUHDNLQJ WKHP L H DEROLVKLQJ WKHLU
SURWHFWLYH IXQFWLRQ RU RU WR SURYH DQG TXDQWLB WKHLU
VHFXULWB FUBSWDQDOBVLV LV WKXV WKH FRXQWHUSDUW WR
FUBSWRJUDSKB ERWK DUH VXEILHOGV RI FUBSWRORJB"""

def calculate_IC(message):
    alphebets = [0 for i in range(26)]
    n_sum = 0
    for ch in message:
        if ch == " " or ch == "\n" or ch == ",":
            continue
        alphebets[ord(ch) - ord("A")] += 1
        n_sum += 1
    f_sum = 0
    for i in range(26):
        f_sum += alphebets[i] * (alphebets[i] - 1)
        #print(chr(ord("A") + i), ": ", alphebets[i], sep = "")
    return (f_sum / (n_sum * (n_sum - 1)))
print("=========================")
print("Answer to the question 3:")
print(f"Index of Coincidence for message 1 is {calculate_IC(message1)}.")
print(f"Index of Coincidence for message 2 is {calculate_IC(message2)}.")
print(f"Index of Coincidence for message 3 is {calculate_IC(message3)}.")
print(f"Index of Coincidence for message 4 is {calculate_IC(message4)}.")
```

以上是我的第三題的code。

我先將原本的訊息依序輸入至message1 ~ message 4，之後透過calculate_IC這個函式以及IC的計算公式將IC的值算出來。

```
Index of Coincidence for message 1 is 0.06422077622409894.
Index of Coincidence for message 2 is 0.06678956585860447.
Index of Coincidence for message 3 is 0.04942544649037796.
Index of Coincidence for message 4 is 0.06422077622409894.
```

以上是我第三題的輸出結果。

---

4. **Given the following ciphertext, please determine if this encrypted message was enciphered using a monoalphabetic or polyalphabetic cipher based on the message's index of coincidence**

Sol.

```
print("Answer to the question 4:")
ciphertext2 = """RHVST TEYSJ KMHUM BBCLC GLKBM HBSJH HDAYC PPWHD UUTAP STJAI
YMXKA OKARN NATNG CVRCH BNGJU EMXWH UERZE RLDMX MASRT LAHRJ
KIILJ BQCTI BVFZW TKBQE OPKEQ OEBMU NUTAK ZOSLD MKXVO YELLX
SGHTT PNROY MORRW BWZKX FFIQJ HVDZZ JGJZY IGYAT KWVIB VDBRM
BNVFC MAXAM CALZE AYAZK HAOAA ETSGZ AAJFX HUEKZ IAKPM FWXTO
EBUGN THMYH FCEKY VRGZA QWAXB RSMSI IWHQM HXRNR XMOEU ALYHN
ACLHF AYDPP JBAHV MXPNF LNWQB WUGOU LGFMO BJGJB PEYVR GZAQW
ANZCL XZSVF BISMB KUOTZ TUWUO WHFIC EBAHR JPCWG CVVEO LSSGN
EFGCC SWHYK BJHMF ONHUE BYDRS NVFMR JRCHB NGJUB TYRUU TYVRG
ZAXWX CSADX YIAKL INGXF FEEST UWIAJ EESFT HAHRT WZGTM CRS
"""
print(f"Index of Coincidence for the ciphertext is {calculate_IC(ciphertext2)}.")
print("It is a polyalphabetic cipher.")
```

以上是我第四題的code。

我先將密文塞至ciphertext2裡面，並利用上一題的calculate_IC函式將IC算出來。

```
Answer to the question 4:
Index of Coincidence for the ciphertext is 0.039780853797483695.
It is a polyalphabetic cipher.
```

以上是輸出結果。

因為IC是0.039左右，所以我認為他與英文IC的0.066~0.068不相同，與講義中提到的random的IC(0.038)較為接近，因此我認為他不是monoalphabetic，而是polyalphabetic加密。

Bonus: Suppose a columnar transposition cipher is not 10 column by 5 row
LLOWA POLNH NHOEG YSOKD NDWNI TUIEE FHMDR IEBYT CWEOH ARRUE.
Please break this message and state your method! If you can provide your own algorithm will be plus

Sol.

```
from collections import Counter
import numpy as np
ciphertext_bonus ="L L O W A P O L N H N H O E G Y S O K D N D W N I T U I E E F H M D R I E B Y T C W E O H A R R U E"
A_1 = ciphertext_bonus.split()

vowels = ["A", "E", "I", "O", "U"]
row = 9
column = int(len(A_1) / row) + 1
for i in range(row * column - len(A_1) - 1, 0, -1):
    A_1.insert(-(column - 1) * i, 0)
A_1.append(0)
A_2 = np.reshape(A_1, (row, -1))
for A_row in A_2.tolist():
    print([ch for ch in A_row if ch != "0"])
```

以上是我bonus的code。
算法的部分是先把題目的字串先切成我猜想大小的矩陣，如果字串不夠的部分後面就會補0，最後在print出來的時候不會把0印出來。

```
['L', 'L', 'O', 'W', 'A', 'P']
['O', 'L', 'N', 'H', 'N', 'H']
['O', 'E', 'G', 'Y', 'S', 'O']
['K', 'D', 'N', 'D', 'W', 'N']
['I', 'T', 'U', 'I', 'E', 'E']
['F', 'H', 'M', 'D', 'R']
['I', 'E', 'B', 'Y', 'T']
['C', 'W', 'E', 'O', 'H']
['A', 'R', 'R', 'U', 'E']
```

以上是我的輸出結果。
The plaintext is LOOK IF I CALLED THE WRONG NUMBER WHY DID YOU ANSWER THE PHONE

因為題目說不是5 x 10的矩陣，所以我就猜想應該是將明文塞進一個比明文更大的矩陣裡面，最後試到 9 rows x 6 columns的時候就成功將密碼破譯了。
不過我的code裡面的rows應該是老師的columns，也就是應該把LLOWAP從上往下放而不是從左往又放，所以我放的方向不同可能會導致答案跟老師的不太相同，像是老師的放法可能會說他是6 rows x 9 columns，但是實際上我解出來的明文也會是對的。