COSC 3371
Spring 2024

Homework Assignment 1

Please complete the attached Python source file (HW1.py) to solve the following problems. For each problem, replace the code **between** `// BEGIN SOLUTION` and `// END SOLUTION` with your solution (please do not modify other parts of the code). The submission uploaded to Canvas should include the completed Python source file. Please ensure that you use only standard libraries and that the uploaded source file can be compiled and executed without errors and unhandled exceptions. In each problem, your goal is to recover a plaintext from a given ciphertext.

Problem 1 (30 points): Substitution Cipher

Agent James Vond,

One of our secret agents, Agent 006, has recently gone missing in the Caribbean. At the time of his disappearance, he was investigating a reclusive billionaire, Dr. On. We do not have any information that would connect Dr. On to criminal activities, but our agent was quite insistent on the investigation. This was the last message that we received from our agent:

"ROYQWH KQXXJYQ: N LQGNQAQ HDJH FO. VW NX J KQKLQO VZ J XQMOQH MONKQ VOYJWNSJHNVW MJGGQF U.D.J.W.H.V.K., IDVXQ YVJG NX HVHJG IVOGF FVKNWJHNVW. HDQNO UGJW NX HV JMBRNOQ J XRUQOIQJUVW JWF HV DVGF HDQ IVOGF OJWXVK. N JK JZOJNF HDJH IQ FV WVH DJAQ KRMD HNKQ LQZVOQ HDQT XRMMQQF.

N DJAQ OQMQWHGT NWHQOMQUHQF JW QWMOTUHQF KQXXJYQ (JHHJMDKQWH MNUDQO2.HCH) HDJH IJX XQWH LT FO. VW HV VWQ VZ DNX MVWXUNOJHVOX, HDQ NWZJKVRX KO. LGVIZNQGF. N KJWJYQF HV FNXMVAQO HDJH HDQ KQXXJYQ IJX QWMOTUHQF RXNWY HDQ PJMEJG MNUDQO (XQQ XVROMQ MVFQ), LRH N IJX WVH JLGQ FNXMVAQO HDQ XQMOQH EQT, JWF HDQ MNUDQO XQQKX HV LQ RWLOQJEJLGQ. N JK JZOJNF HDJH FQMOTUHNWY HDNX KQXXJYQ NX HDQ VWGT IJT HV XHVU FO. VW'X VOYJWNSJHNVW.

UGQJXQ XQWF OQNWZVOMQKQWHX NKKQFNJHQGT! N HONQF HV JMH MJRHNVRXGT, LRH N DJAQ J ZQQGNWY HDJH FO. VW'X DQWMDKQW JOQ VWHV KQ. N FVW'H EWVI DVI GVWY N DJAQ LQZVOQ HDQT FNXMVAQO KT OQJG NFQWHNHT JWF KT XQMOQH DNFNWY UGJ"

We believe that the message was encrypted using a substitution cipher, but we do not have the key to decrypt it. Agent Vond, we task you with decrypting the message and finishing the investigation. Since Agent 006 disappeared without a trace under such suspicious circumstances, it is imperative that you discover what happened as soon as possible.

Sincerely,

M

The ciphertext was encrypted using a substitution cipher, and the plaintext is an English-language text. Note that whitespaces and punctuations are not encrypted.

- Compute and print the frequency of each letter (from `A` to `Z`) in the ciphertext (15 points).
- Decrypt the ciphertext and print the plaintext (15 points). You can manually identify which plain letter is substituted for which cipher letter by comparing the computed frequencies with the following typical frequencies:

E:	0.108	D:	0.034	B:	0.013
T:	0.075	C:	0.032	F:	0.011
A:	0.067	M:	0.027	V:	0.008
O:	0.058	L:	0.025	K:	0.004
I:	0.055	P:	0.016	Z:	0.002
N:	0.051	Y:	0.016	J:	0.001
R:	0.047	G:	0.015	Q:	0.001
S:	0.047	U:	0.014	X:	0.001
H:	0.037	W:	0.013		

The frequencies of letters in the homework may not match this table entirely, but they should not differ too much.

To decode the message, you should calculate the frequencies of letters in the ciphertext and compare them with the probability distribution above, using integers instead of probabilities. In addition to the frequencies, you may use other knowledge about English words to help decrypt. Write a report on your experience of the process, such as some of the techniques you used.

You may want to develop a short program that helps you go through the decoding process.

The output of this part includes:

- A frequency table for the ciphertext.
- The decrypted message. The cyphertext was given in UPPER cases. I suggest that you use lowercase letters for the plaintext so you can know the progress of decryption.

Read the decrypted message to give you more information on what to do next.

Problem 2 (20 points): Brute Force

Once you have decrypted Agent 006's message, you realize that the fate of the world is at stake. Unfortunately, the design of the Jackal cipher is somewhat confusing, and there is no time to analyze it, given the urgency of the situation.

Notice that the number of possible keys for the Jackal cipher is very low, and we can assume that the plaintext is English-language text. Can you find the correct key and decrypt the ciphertext? Hint: You do not need to understand the cipher or its implementation to find the key.

The program template includes a JACKAL_Decrypt () function. You must understand how to decrypt the second message using this function, but you do not need to understand the detailed steps of the function. The function uses two integers to encrypt or decrypt a message. Unfortunately, we don't know the key (the two integers between 0 and 127). That is where you will use brute force.

The output of Problem 2 is the plaintext of the second message. Please read it. It contains information on how to decrypt the third message.

Problem 3 (20 points): "One-Time" Pad

After decrypting the message, you immediately fly to Hawaii and prepare to intercept the exchange. With the element of surprise on your side, you easily defeat the agents of P.H.A.N.T.O.M. You expect to retrieve the secret plans from them; unfortunately, all you find is a USB drive with a single encrypted file.

You have everything you need for decrypting cipher3.txt: the cipher algorithm ("onetime" pad with repeating key) and the secret key. Hint: the binary XOR operation in Python can be performed using the ^ operator.

The output is the decrypted message.