Assignment 2

Problem 1-2. One-time pad with ciphertext feedback

It is well known that re-using a "one-time pad" can be insecure. This problem explores this issue, with some variations.

In this problem all characters are represented as 8-bit bytes with the usual US-ASCII encoding (e.g."A" is encoded as 0x41). The bitwise exclusive-or of two bytes x and y is denoted $x \oplus y$.

Let $M = (m_1,m_2,..., m_n)$ be a message, consisting of a sequence of n message bytes, to be encrypted. Let $P =(p_1,p_2,...,p_n)$ denote a pad, consisting of a corresponding sequence of (randomly chosen) "pad bytes" (key bytes).

In the usual one-time pad, the sequence $C =(c_1,c_2,...,c_n)$ of ciphertext bytes is obtained by xor-ing each message byte with the corresponding pad byte:

$c_i = m_i \oplus p_i$, for i =1 ...n .

When we talk about more than one message, we will denote the messages as M1,M2,...,Mk and the bytes of message M_j as m_{ji}, namely $Mj =(m_{j1},...,m_{jn})$; we'll also use similar notation for the corresponding ciphertexts.

(a) Here are two 8-character English words encrypted with the same "one-time pad". What are the words?

e9 3a e9 c5 fc 73 55 d5

f4 3a fe c7 e1 68 4a df

Describe how you figured out the words.

(b) Ben Bitdiddle decided to fix this problem by making sure that you can't just "cancel" pad bytes by xor-ing the ciphertext bytes. In his scheme the key is still as long as the ciphertext. If we define c0 = 0 for notational convenience, then the ciphertext bytes c_1, c_2,...,c_n are obtained as follows:

$$c_i = m_i \oplus ((p_i + c_i - 1) \bmod 256) .$$

That is, each ciphertext byte is added to the next key byte and the addition result (modulo 256) is used to encrypt to the next plaintext byte.

Ben is now confident he can reuse his pad, since $(k_i + c_i - 1) \bmod 256$ will be different for different messages, so nobody would be able to cancel the k_i's out. You are provided with otp-feedback.py, which contains an implementation of Ben's algorithm.

You are also given the file tenciphs.txt, containing ten ciphertexts $C_1, C_2, ..., C_{10}$ produced by Ben, using the same pad P . You know that these messages contain valid English text.

Submit the messages and the pad, along with a careful explanation of how you found them, and any code you used to help find the messages. The most important part is the explanation.