Q4. Protein Translation (30 marks):

Ribonucleic Acid (RNA) is the genetic sequence used to create proteins and it is translated from Deoxyribonucleic Acid (DNA) sequence.

A DNA sequence comprises of the combination of the four nucleotide bases: A, C, G, and T. During RNA translation, the base 'T' in the DNA sequence is translated to 'U', and the rest of the bases remain unchanged.

After the RNA sequence is obtained, it is translated into codons. A codon is a sub-sequence which contains three nucleotides. These codons are used to create amino acids for protein synthesis. Such protein synthesis starts with methionine "AUG" (which is also known as the start codon) and terminates with one of these three stop codons: UAG, UAA, or UGA. That is, a complete protein sequence should begin with a start codon and end with a stop codon.

Write a program to:

Input, one DNA fragment with not more than 100 characters. The input characters can be in capital or small letters.

Output, according to the following instructions.

(1) If the input DNA fragment cannot form a complete protein sequence, then output "0".

(2) If the input DNA fragment can form a complete protein sequence, then output the following in sequence:
   a. The RNA sequence translated from the above DNA fragment.
   b. The position of the start codon in the RNA sequence.
   c. The total number of founded codons, EXCLUDING the stop codon.
   d. The founded codons (EXCULDING the stop codon) in the sequence how they appear in the RNA sequence. There must be a space between two neighbouring codons.

(**Note**: the output characters **must** be all in capital letters.)

试题 4. 蛋白质的翻译 （30 分）：

核糖核酸 （RNA）是用于产生蛋白质的基因序列，它是从脱氧核糖核酸（DNA）序列翻译而来的。

DNA 序列由四个核苷酸碱基 (A，C，G 和 T) 所组成。在 RNA 的翻译过程中，DNA 序列中的碱基 "T" 将被翻译成 "U"，其余的碱基则保留不变。

通过以上过程所得到的 RNA 序列将可进一步分解成密码子。每个密码子是由三个核苷酸碱基所构建。这些密码子将用于产生合成蛋白质的氨基酸。此蛋白质合成的程序将从蛋氨酸 "AUG" 开始（这也被称为起始密码子），并以以下三个终止密码子之一作为结束：UAG，UAA 或 UGA。换句话说，完整的蛋白质序列应以起始密码子开头，并以终止密码子结束。

试写一程式以

输入 一个 DNA 片段，已知此片段包含不超过 100 个字母。输入的字母可以是大写或小写。

根据以下规定，输出

(1) 如果输入的 DNA 片段不能形成完整的蛋白质序列，则输出"0"。

(2) 如果输入的 DNA 片段可以形成完整的蛋白质序列，则依次输出：

a. 从上述 DNA 片段翻译而来的 RNA 序列。
b. 此 RNA 序列中起始密码子的位置。
c. 此 RNA 序列中密码子总数，**不包括**终止密码子。
d. 按照它们在 RNA 序列中出现顺序，列出所有找到的密码子（但**不包括**终止密码子）。两个相邻密码子之间必须有一个空格。

（**注意**：输出字母**必须**全部为大写字母。）

Example (例子)

Input (輸入)	Output (輸出)
TGAGAATCGA	0
accgtaatgc	0
ggaccatgaccaccatttaaa	GGACCAUGACCACCAUUUAAA 6 4 AUG ACC ACC AUU
atggcttatggatagtacttctccataga	AUGGCUUAUGGAUAGUACUUCUCCAUAGA 1 4 AUG GCU UAU GGA
gttttcccctcccatgtgctagttgctttatctgttcact	GUUUUCCCCUCCCAUGUGCUAGUUGCUUUAUCUGUUCACU 14 2 AUG UGC