Gene enumeration   X48860_en

Recall that in prokaryotic genomes, the sequence coding for a protein occurs as one contiguous open reading frame, and that an open reading frame begins with the start codon ATG (methionine) in most species and ends with a stop codon TAA, TAG, or TGA.

For example, the DNA sequence of Bacteriophage ϕ-X174, which was the first genome to be sequenced, has 117 open reading frames (11 of which are protein coding genes) within a circular single strand of 5,386 nucleotides.

Write code for the gene enumeration problem. The program must implement and use the GENE-ENUMERATION function in the pseudocode discussed in class, which is iterative and is not allowed to perform input/output operations. Make one submission with Python code and another submission with C++ code.

Input

The input is a string s over the alphabet $\Sigma = \{A, C, G, T\}$.

Output

The output is all minimal substrings of s (all open reading frames) from a start codon to a stop codon.

Sample input

```
GAGTTTTATCGCTTCCATGACGCAGAAGTTAACACTTTCGGATATTTCTGATGAGTCGAAAAATTATCTTGATAAAGCAGGAATTACTACTGCTTG
```

Sample output

```
ATGACGCAGAAGTTAACACTTTCGGATATTTCTGATGAGTCGAAAAATTATCTTGATAAAGCAGGAATTACTACTGCTTGTTTACGAATTAAATCG
ATGAGTCGAAAAATTATCTTGATAAAGCAGGAATTACTACTGCTTGTTTACGAATTAAATCGAAGTGGACTGCTGGCGGAAAATGAGAAAATTCGA
ATGAGAAAATTCGACCTATCCTTGCGCAGCTCGAGAAGCTCTTACTTTGCGACCTTTCGCCATCAACTAACGATTCTGTCAAAAACTGACGCGTTG
ATGAGGAGAAGTGGCTTAATATGCTTGGCACGTTCGTCAAGGACTGGTTTAGATATGAGTCACATTTTGTTCATGGTAGAGATTCTCTTGTTGACA
ATGCTTGGCACGTTCGTCAAGGACTGGTTTAGATATGAGTCACATTTTGTTCATGGTAGAGATTCTCTTGTTGACATTTTAAAAGAGCGTGGATTA
ATGAGTCACATTTTGTTCATGGTAGAGATTCTCTTGTTGACATTTTAA
ATGGTAGAGATTCTCTTGTTGACATTTTAA
ATGCTGTTCAACCACTAA
```

Problem information

Author : Gabriel Valiente
Generation : 2022-07-07 18:26:44