UCSC Silicon Valley Extension

C Programming, Advanced

Assignment 1 : C Review
Instructor : Radhika Grover

Include test cases in all of the following programs to demonstrate that the program executes correctly. The output carries 50% of the points for that problem. No points will be awarded if the program does not compile.

Problem 3 "It's all in the genes"
DNA forms the building blocks for the each cell in the human body. DNA creates RNA which is made up of four bases. A, C, G and U. The bases can be combined together to create a triplet known as a codon that is responsible for protein synthesis. The following are examples of codons:
UAA, CGC, CGA, CGG, GUA, AUG, AUC, UAA, UAG, GGU, GUA

Any one of the four RNA bases may be present in one of three positions in the codon, leading to 64 possible codon combinations. The genetic code is comprised of the set of all possible codons. Each RNA sequence consists of a set of codons such as AUG, CUA, UUC, and UAA. The four codons AUG, UAA, UAG, and UGA have a special meaning: AUG indicates start of the RNA sequence, and UAA, UAG, UGA indicates the end of the sequence.

DRTAYUIOHN
TGHUIIOPKL
BAUGUUCUGA
ERTUSDFNMU
RTYAIOPKLI
GHCAYUIOXC

The vertical sequence, AUGUAA contains 2 codons. The horizontal sequence AUGUUCUGA has 3 codons. In this problem, you have to find the RNA sequence with the largest number of codons. The sequences are hidden in a 2D grid in the horizontal and vertical directions.

An input file called codons.txt containing the names of all the codons is provided below:

64 # number of codons
UUU

UUC
UUA
UUG
CUU
CUC
CUA
CUG
AUG
GUU
GUC
GUA
GUG
UCU
UCC
UCA
UCG
AGU
AGC
CCU
CCC
CCA
CCG
ACU
ACC
ACA
ACG
GCU
GCC
GCA
GCG
UAU
UAC
CAU
CAC
CAA
CAG
AAU
AAC
AAA
AAG
GAU
GAC
GAA

GAG
UGU
UGC
UGG
CGU
CGC
CGA
CGG
AGA
AGG
GGU
GGC
GGA
GGG
AUU
AUC
AUA
UAA
UAG
UGA

The test case file contains data in the format
10 #Number of test cases
9 15 #Number of rows =9 Number of columns = 15

Your program should read the test cases from the input file provided below and report the output in the following format:
Test case 1: Maximum number of codons: #
Test case 2: Maximum number of codons: #
and so on

Reference for codons: https://en.wikipedia.org/wiki/Genetic_code