## Jutge.org

The Virtual Learning Environment for Computer Programming

Words 2 X86108\_en

Nucleic acid sequences are labeled over the alphabet  $\{A, C, G, T\}$ , and there are  $4^n$  possible genomic sequences of length n. Amino acid sequences, on the other hand, are labeled over the alphabet  $\{A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y\}$ , and there are  $20^n$  possible proteomic sequences of length n. An interesting problem is the generation of all the genomic sequences with n nucleotides or all the proteomic sequences with n amino acids, that is, the generation of all the words of length n over an alphabet  $\Sigma$ .

Write pseudocode, Python code, and C++ code for the words problem. The program must implement and use the words function in the pseudocode, which must be recursive and is not allowed to perform input/output operations. Make two submissions, including the pseudocode as a comment to both the Python and the C++ code.

## Input

The input is an integer n and an alphabet  $\Sigma$ .

## Output

The output is a sorted list of all the words of length n over the alphabet  $\Sigma$ .

Sample input 1	Sample output 1
1 G T A C	A C G T
Sample input 2	Sample output 2
2	AA

Sample input 2	Sa
2	AA
G T A C	AC
	AG
	ΑT
	CA
	CA CC CG
	CG
	СТ
	GΑ
	GC
	GG
	GΤ
	TA
	TC
	ΤG

Sample input 3	Sample output 3
3	AAA
G T A C	AAC
	AAG
	AAA AAC AAG AAT

TT

ACA ACC ACG ACT AGA AGC AGG AGT ATA ATC ATG ATT CAA CAC CAG CAT CCA CCC CCG CCT CGA CGC CGG CGT CTA CTC CTG CTTGAA GAC

GAG GAT GCA GCC GCG GCT GGA GGC GGG GGT GTA GTC GTG GTT TAA TAC TAG TAT TCA TCC TCG TCT TGA TGC TGG TGT TTA TTC TTG TTT

## **Problem information**

Author: Gabriel Valiente

Generation: 2021-11-20 09:06:43

© *Jutge.org*, 2006–2021. https://jutge.org