Read abundance X28783_en

The abundance (number of occurrences) of a read in a read set is an indicator value for read confidence in high-throughput sequencing studies.

Write pseudocode, Python code, and C++ code for the read abundance problem. Make two submissions, including the pseudocode as a comment to both the Python and the C++ code.

Input

The input is a collection of n strings $S = \{s_1, s_2, \dots, s_n\}$ (genomic sequence reads, possibly reverse complemented) over the alphabet $\{A, C, G, T\}$.

Output

The output is the sorted frequency distribution of S.

Sample input 1

```
TCATC
TTGAT
TCATC
TGAAA
GATGA
TTTCA
ATCAA
TTGAT
TTTCA
```

Sample output 1

```
ATCAA 3
GATGA 3
TGAAA 3
```

Problem information

Author: Gabriel Valiente

Generation: 2026-01-25T14:54:31.664Z