

2E Implement GreedyMotifSearch with Pseudocounts

Greedy Motif Search with Pseudocounts Problem

Implement *GreedyMotifSearch* with pseudocounts.

Input: A collection of strings *Dna*, and integers *k* and *t*.

Output: A collection of strings resulting from running *GREEDYMOTIFSEARCH*(*Dna*, *k*, *t*) with pseudocounts.

| | | | | | |
|---|------|------|------|------|------|
| A | 6/10 | 1/10 | 2/10 | 0/10 | 1/10 |
| C | 2/10 | 1/10 | 4/10 | 0/10 | 7/10 |
| G | 0/10 | 0/10 | 2/10 | 2/10 | 0/10 |
| T | 2/10 | 8/10 | 2/10 | 8/10 | 2/10 |

| | | | | | |
|---|------|------|------|------|------|
| A | 7/14 | 2/14 | 3/14 | 1/14 | 2/14 |
| C | 3/14 | 2/14 | 5/14 | 1/14 | 8/14 |
| G | 1/14 | 1/14 | 3/14 | 3/14 | 1/14 |
| T | 3/14 | 9/14 | 3/14 | 9/14 | 3/14 |

Formatting

Input: Space-separated integers *k* and *t*, followed by a newline-separated collection of strings *Dna*.

Output: A space-separated list of strings containing a collection of strings resulting from running *GREEDYMOTIFSEARCH*(*Dna*, *k*, *t*) with pseudocounts (If at any step you find more than one *Profile*-most probable *k*-mer in a given string, use the one occurring first).

Constraints

- The integer *k* will be between 1 and 10^2 .
- The integer *t* will be between 1 and 10^2 .
- The number of strings in *Dna* will be between 1 and 10^2 .
- The length of each string in *Dna* will be between 1 and 10^2 .
- Each string in *Dna* will be a DNA string.

Test Cases

Case 1

Description: The sample dataset is not actually run on your code.

Input:

3 5

GGCGTTC GGC G TC GTC C GG GTTCGC C CGTC TC C C T T TTCG

Output:

TTC TC TTC TC TTC

Case 2

Description: A larger dataset of the same size as that provided by the randomized autograder. Check input/output folders for this dataset.