

1I Find the Most Frequent Words with Mismatches in a String

Frequent Words with Mismatches Problem

Find the most frequent k -mers with mismatches in a string.

Input: A DNA string $Text$ as well as integers k and d .

Output: All most frequent k -mers with up to d mismatches in $Text$.

CGACTAGTTT
ATT ATT
ATT

Formatting

Input: A DNA string $Text$ as well as integers k and d .

Output: A space-separated list of strings representing all most frequent k -mers with up to d mismatches in $Text$.

Constraints

- The length of $Text$ will be between 1 and 10^3 .
- The integer k will be between 1 and 10^1 .
- The integer d will be between 1 and 10^1 .
- $Text$ will be a DNA string.

Test Cases

Case 1

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

Input:

```
ACGTTGCATGTCGCATGATGCATGAGAGCT
4 1
```

Output:

```
ATGC ATGT GATG
```

Case 2

Description: *Text* contains partial and complete matches for the most frequent word.

Input:

```
AGGT
2 1
```

Output:

```
GG
```

Case 3

Description: $d = 0$.

Input:

```
AGGGT
2 0
```

Output:

```
GG
```

Case 4

Description: *Text* has multiple most frequent words.

Input:

```
AGGCGG
3 0
```

Output:

```
AGG GGC GCG CGG
```

Case 5

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