

BioTech DNA Sequencer v1.1

Your task is to create an application that would read in a DNA sequence and allow our scientists to run some analysis on it. A sample text file has been provided for you (see attached). A DNA sequence contains a series of nucleotides represented as letters -- Letters G, A, C or T is used to identify the nucleotides in the the sequence. For each nucleotide, there is also a number associated with it that represents the strength of its bonding. This number ranges from 1 to 5. Make sure that your program reads the DNA sequence from the provided text file at the very beginning [30] and provides the following features to the scientists (users of the program).

[50] Feature #1 Nucleotide Report:

Display how many times each nucleotide appears in the sequence regardless of its bonding strength. So when end user chooses this feature, you should display how many times A appears in the sequence, how many times G appears in the sequence, how many times C appears in the sequence, and how many times T appears in the sequence.

```
Welcome to BioTech DNA Sequencer v1.1
1- Nucleotide Report
2- Pattern Search
3- Strength Filtering
4- Avg. Sequence Strength
0- Exit
1
# of G: 13
# of A: 13
# of C: 8
# of T: 12

Welcome to BioTech DNA Sequencer v1.1
1- Nucleotide Report
2- Pattern Search
3- Strength Filtering
4- Avg. Sequence Strength
0- Exit
2
Enter the first letter of the pattern: T
Enter the second letter of the pattern: G
Enter the third letter of the pattern: T
We found the pattern in the DNA sequence at index 39.

Welcome to BioTech DNA Sequencer v1.1
1- Nucleotide Report
2- Pattern Search
3- Strength Filtering
4- Avg. Sequence Strength
0- Exit
3
Enter the minimum required strength: 4
Here is the sequence:
TGACCATAGTG

Welcome to BioTech DNA Sequencer v1.1
1- Nucleotide Report
2- Pattern Search
3- Strength Filtering
4- Avg. Sequence Strength
0- Exit
4
The Avg. Strength of the DNA sequence is: 2
```

[50] Feature #2 Pattern Search:

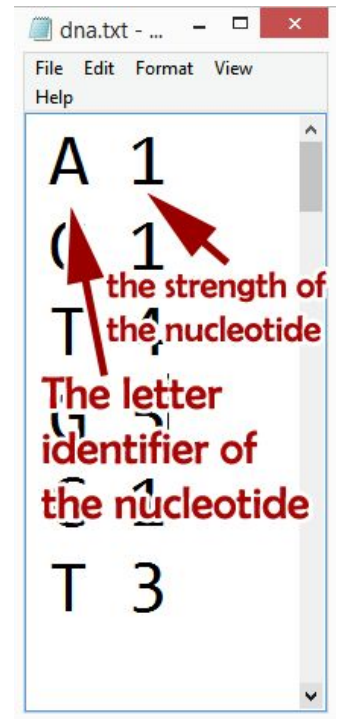
Ask the end user to enter three letters (letters can be either G, A, C or T) -- these three letter sequence is called a pattern. Check if within the sequence, this three letter pattern appears back to back. For example, if user enters A C A, then check the DNA sequence if there is a pattern A C A appear in it. If it does, tell the user the index (zero based index) at which the pattern occurs. Otherwise, inform the user that the pattern cannot be found.

[50] Feature #3 Strength Filtering:

Ask the end user to enter the minimum strength (a number between 1 and 5 inclusive) and display the full DNA sequence on screen excluding the nucleotides with strength less than the entered minimum strength.

[50] Feature #4 Avg. Sequence Strength:

Display the average strength of the whole DNA sequence. That is the sum of strength of each nucleotide in the sequence divided by the number of nucleotides in the sequence.



[20] Exit: Allow user to exit the program.

Make sure to create a class named Nucleotide, the class should have two fields: 1- *ID* (the letter designation of type string) and 2- *Strength* (the bonding strength of type integer). You need read the file into an array of Nucleotide class and use this array to perform all the features.

Avoid Penalties!

-100 points penalty if your code is not properly indented or contains no comments.

-150 points penalty if you do not use functions, class or your code does not compile or you use global variables.