

CIS 1.5 Fall 2009 Homework IV

Overview

- This is the fourth homework assignment for CIS 1.5.
- The entire assignment will be worth 10 points, that is 10% of your semester grade.
- **It is due by midnight on Friday November 13th** and must be submitted by email (as below).
- **Follow these emailing instructions:**
 1. Create a mail message addressed to *parsons@sci.brooklyn.cuny.edu* with the subject line **CIS 1.5 HW4**.
 2. Write your name, that is the name under which you registered for the course, in the email (when I get an email from *deathmetal@aol.com* or *pinkprincess@yahoo.com*, I can usually guess whose program it is, but that is not as good as *knowing* whose program it is).
 3. Attach **ONLY** the **.cpp** files as I tell you below.
 4. Use a **zip** utility to bundle all your files together and send them as ONE attachment to the email.
on a PC: use **WinZip**
on a Mac: use **File - Create Archive...**
on Linux: use **zip**
 5. Failure to follow these instructions will result in points being taken away from your grade. The number of points will be in proportion to the extent to which you did not follow instructions. . . (which can make it a lot harder for me to grade your work — grrrr!)

Description

For this project, you will manipulate some strings that represent lengths of DNA, writing programs to simulate some of the common things that one does with DNA sequences in bioinformatics applications.

1. Write a program **bio.cpp** that does the following:
 - First, the program should read two strings in from the file `dna.txt` into the variables `dna1` and `dna2`. You can find `dna.txt` on the class website (it is on the page for Unit IV).
 - Then the program should prints the two strings from the file in reverse.

(2 points)
2. A function to test part of a string
 - Now extend your program. Add a function `isItT` which tests if a character is the letter `t`. That is, the function should take a character as an argument, and return true if that character is a `t` and false otherwise.
 - Use `isItT` to print:

```
The third character is a t
```

if the third character of `dna2` is a `t`

(2 points)
3. Calculate the complement of a string

As mentioned above, the two strings you read from `dna.txt` are representations of short lengths of DNA, with the letters `t`, `a`, `c` and `g` representing the four building blocks from which DNA is constructed. These building blocks have special properties, among which are the fact that `a` is the *complement* of `t`, and `g` is the complement of `c`.

- Use this fact to write a function `complement` that computes the complement of each of the four letters in a string. This is a function that takes a character as a parameter, and returns a character. Thus:
`complement(c)` will return `g`
`complement(t)` will return `a`
and so on.
- Now, get your program to concatenate `dna1` and `dna2`, and then step through the string character by character printing out the complement of each character in the string.

(2 points)

4. A function that complements

- Now write a function called `complementary` which takes as its argument a string, and returns the complement of that string.
- `complementary` can do this by using the function `complement` that you wrote for the previous question — `complementary` can call `complement` on each element of the string in turn.
- Your program should call `complementary` to compute the complement of `dna1`, and assign this value to the string `dna3`.
- Your program should then call `complementary` to compute the complement of `dna2`, and assign this value to the string `dna4`.
- Your program should print the result of concatenating `dna3` and `dna4`.

(2 points)

5. A function that counts

- Finally, write a function `countTheTs` that takes as its argument a string, and returns an integer that gives the number of `ts` in that string.
- Your program should use `countTheTs` to calculate the number of `ts` in the complements of `dna1` and `dna2`
- Print out this number.

(2 points)

Submission

- You will be submitting one file: **bio.cpp**
- Make sure that you have a comment at the top of the file that contains the name of the file, your name, “CIS 1.5 HW4” and the submission date (November 13, 2009).
- The subject line of your email should say: CIS 1.5 HW4
- The body of your email should contain your name.