

CS 150: Intro to OOP, Spring 2012

Assignment 3

Due *via email* by 11:59pm on Monday, February 13, 2012

For this assignment, you may work in pairs to complete the work; in fact, I encourage all of you to find a partner for the assignment. If you would like a partner but are unable to find one, please contact the instructor to be paired up with another interested student.

Exactly one member of the team should email the program file to the instructor at echambe5-at-slu.edu by 11:59pm on the date due. Please remember to include the name of the other person who worked on the program in the email.

DNA can be modeled as a string of characters using the alphabet **A**, **C**, **G**, and **T**. One form of DNA mutation occurs when a substring of the DNA is reversed during the replication process. Usually, such a reversal occurs between what are termed *inverted pairs*, where some substring is followed later by its reversal. For example, if the pattern **TGAA** is later followed by the inverted pattern **AAGT**, it is possible that the slice of DNA delimited by those patterns could be inverted and reattached, since the bonds at each end will be locally the same. As an example, the string **TGAACATTAAGT** would mutate to **TGAATTACAAGT** - the middle 4 characters are flipped, but the inverted pair is left the same.

Design a program that works as follows. It should ask the user for an original DNA string as well as the particular pattern that is inverted. It should then locate the leftmost occurrence of that pattern, and the next subsequent occurrence of the inverted pattern. The output should be the mutated DNA, with the segment between the inverted pair reversed. An example session might proceed as follows (where the user input is anything past the colons in the two first lines):

```
Enter a DNA sequence: CGATTGAACATTAAGTCCATT
```

```
Enter the pattern: TGAA
```

```
Mutated DNA sequence: CGATTGAATTACAAGTCCAATT
```

Please note the following things to be careful of in your program:

- Your program should work for *any* string and pattern that is entered as input. So don't only solve this for the example and expect it to work - you'll need to try entering a variety of inputs to be sure it is actually going to work on all the inputs I could possibly try!
- Be sure to test a variety of cases. Does your code work if there are multiple copies of the pattern or the reverse? How about if the reverse comes before the pattern? What if the two overlap? Any of these cases could crash your code if you aren't careful, so be sure to test them before handing in.
- You should only replace the *first* substring which is between the pattern and its reverse, not all of them (if the pattern and reverse happen to appear many times).
- As always, you are expected to include comments and use meaningful variable names in your program.

- Any code which is submitted that does not at least compile - i.e. I can type “python program3.py” and enter input strings without getting errors - will immediately lose 20% of the points.

Also, I strongly recommend that you reread section 2.3.1 on strings before beginning this assignment, in order to refamiliarize yourselves with the available methods for this data type.