Programming Exercise: Strings in Java

Assignment 1: Processing DNA Strings

Modify the program you wrote in the last lesson to also be able to answer the following questions. Make sure the program works regardless of the case of the string. Be sure to test your program on a small set of data first.

- How many times does the codon CTG appear in a strand of DNA?
- What is the length of the longest gene in a collection of genes (that is, the longest String in a StorageResource of strings)?

Assignment 2: Find All URLs

Write a program to find all the URLs on a web page and to

- print the URLs,
- print the number of URLs found,
- print the the number of secure links found,
- print the number of links that have ".com" in them,
- print the number of links that end with ".com" or ".com/" and
- print the total number of "." that appear in all the links.

More specifically, you should:

- Write the method `findURLs` that has one String parameter `url`, and it returns a StorageResource of strings that are all the URLs on the web page read into `url`. A URL on a page comes immediately after `href=` and is in quotes, for example `href="https://amazon.com"`. One can find URLs on a page by searching for "href=" to identify the start of a URL, and then searching past "href=" to find the next quotation marks.
- Write the void method `testURLWithStorage`, which has no parameters. This method should do the following.

- ○ Call `findURLs` with the URL address:

  http://www.dukelearntoprogram.com/course2/data/manylinks.html

- ○ Print all the URLs found on this page

- ○ Print the number of URLs found on this page

- ○ Print the number of secure links found on this page (those URLs that start with https)

- ○ Print the number of links that have ".com" in the link

- ○ Print the number of links that end with ".com" or ".com/"

- ○ Print the total number of "."'s that appear in all the links. (note the dot may appear more than once in a link such as www.amazon.com)