**Bioinformatics 529 Homework 1**

**Amelia Lauth**

**01/27/22**

**Resources used (Including websites, partners in class, etc.):**

Partners in Class: Crystal Wen, Jianhui Gong, Matthew Pun, Noah Helton

Brad Crone, Tutor Catherine Barnier

BIOINF 575 past HW (especially with the Counter), BIOINF 529 Class\_06 solutions (with the De Bruijn Class)

<https://www.w3schools.com/python/ref_func_zip.asp>

<https://docs.python.org/3/library/random.html>

<https://docs.python.org/3/library/collections.html>

<https://moonbooks.org/Articles/How-to-select-randomly-keys-from-a-dictionary-in-python-3-/>

<https://www.w3schools.com/python/ref_string_replace.asp>

<https://www.geeksforgeeks.org/python-get-dictionary-keys-as-a-list/>

<https://www.w3schools.com/python/ref_string_join.asp>

<https://www.w3schools.com/python/gloss_python_string_concatenation.asp>

<https://stackoverflow.com/questions/45340785/update-counter-collection-in-python-with-string-not-letter>

<https://www.geeksforgeeks.org/python-random-sample-function/>

**Statement of Objective (What was the purpose of this homework assignment):**

The purpose of this homework assignment is to build a large function, incorporating several smaller functions and a class/objects, to read in sequences from a fasta file, splitting the sequences into k-length k-mers, correcting them for sequencing errors in a de Bruijn assembly, building the de Bruijn graph using our de Bruijn class, and finally outputting a Eulerian walk from that graph.

**Procedure (Explain in general terms how you went about implementing the homework assignment):**

I began working through one function at a time, starting with the first function given in the homework, and building up to the last function that incorporated all the previous functions. As I built each function, I performed multiple sanity checks with different scenario inputs to see if my outputs made sense before moving onto the next function. I also discussed in detail with classmates and my tutor what the functions were asking, talking/drawing out diagrams of what needed to be done before coding. I also reached out to Brad for clarification on several functions.

**Difficulties and Roadblocks (What were the pain points in the implementation of this homework assignment):**

Biggest roadblock for me was understanding the pseudocode, the docstrings, and ultimately understanding what the problem/functions were each asking. I found the problems were not clearly phrased, so I struggled on how to approach the problems when coding. To get as much clarification as I could, I reached out to Brad and fellow classmates, who either already understood what the function was asking or had asked Alan and Ryan already. My tutor also was very helpful in looking at the material, understanding what the functions were asking, and explaining it to me. Once I understood what was being asked, I could think of several approaches code-wise to use to solve the problems.