```
/* CMPUT 606 Exercises                                                    */
/* Dues:                                                                  */
/* #1:    11:55pm, January 29, 2020;                                      */
/* #2:    11:55pm, February 5, 2020;                                      */
/* #3:    11:55pm, February 12, 2020                                      */
/* #4:    11:55pm, February 19, 2020;                                     */
/* #5:    11:55pm, February 26, 2020;                                     */
/* #6:    11:55pm, March 18, 2020                                         */
/* #7:    11:55pm, March 25, 2020                                         */
```

Regardless of the collaboration method allowed, you must always properly acknowledge the sources you used and people you worked with. Your professor reserves the right to give you an exam (oral, written, or both) to determine the degree that you participated in the making of the deliverable, and how well you understand what was submitted. For example, you may be asked to explain any solution that was submitted and why you choose to write it that way. This may impact the mark that you receive for the deliverable.

So, whenever you submit a deliverable, especially if you collaborate, you should be prepared for an individual inspection/walkthrough in which you explain what every line of assignment does and why you choose to write it that way.

```
/* CMPUT 606 Exercise #1                                                    */
/* Due:    11:55pm, January 29, 2020;                                       */
```

In this exercise we reproduce an intuitive shot-gun sequencing project. The following list contains the specifications (10 marks in total):

1. Go to NCBI GenBank (or any other whole genome repository) to download a portion of a whole genome of 10,000 bases.

2. By starving each type of nucleotides, collect a set of reads of lengths no greater than 200.

3. For each read, apply a probability of 0.01 at every position to either mutate the nucleotide, or delete the nucleotide, or insert a random nucleotide, or to terminate the read.

   This revises the read into a new one, which replaces the original read in the set.

4. Apply any algorithm or use any program you might think of to recover the whole genome from the set of reads.

5. If the original whole genome is not recovered correctly, then repeat the above process and report how many times you have repeated to recover the original whole genome nearly correct (say for example, 99% in both length and the nucleotides).

6. Briefly discuss your insights on success and/or non-success.

//End of description of Exercise #1.