

# COP 290- Major Exam (September 2020)

**We will not answer any technical questions. Make reasonable assumptions and mention that in the Readme**

## **Test Duration**

You will have a submission window of 4 hours and have to submit the zipped file on Moodle. The exam will open at **8:30AM** on 8th of September and close at **12:30PM**.

## **Submission Guidelines**

1. You **must** use the lexical analysis tools (lex/yacc or flex/bison) and GNUplot that you have worked with, in earlier assignments.
2. You must have a project directory with the required MAKEFILE as well as a README document.
3. All variable names (except possibly the iteration counters used in the loop ) in your code must have the last two digits of your entry number e.g if your entry number is 2018CS11090, a variable **sum** must be declared as **sum90**.
4. Your MAKEFILE should take your entry number as a parameter. *THIS IS A MUST FOR THE GRADING TO HAPPEN SEMI-AUTOMATICALLY.*
5. Attach a screenshot of the code running in your machine as a proof of your execution that will be used in case the code runs into problems at our end.
6. You will also provide a 2-page document (with a flowchart/diagram of your program) where you will outline the design effort for Part d. This has to be submitted after you have done a final submit on Moodle and there would be another link for that. This is a separate document from your README file.
7. Upload the tar/zip of the project in the first link.
8. At 12:15 PM, you will execute the following commands in the project directory (in this sequence), in a terminal
  - clear
  - date
  - ls -lt

Then you will take a screenshot of the terminal and save the image in the project directory (call this imageProject.png).

9. Once you create the zip file of project directory, you will also save the screenshot of the timestamp of the zip file (say project.zip) by executing `ls -l project.zip` and taking a screenshot (call this imageZip.png).
10. Upload the two png files as a zip in the second link.

# Problem Statement

Your program should take as input the names of 2 text documents. Write a program which lists the information for each of those documents-

- The number of single-letter words in each document (e.g. 'l', 'a') (#a1,#a2)  
(2 marks)
- The number of 3 letter words (e.g. 'put', 'let') (#b1,#b2)  
(3 marks)
- The number of words that contain say upto 3 letters from your first name (any permutation is valid). For example- If your name is 'Daniel Radcliffe', then find the number of words having any 3 letters from the set ['d', 'a', 'n', 'i', 'e', 'l'] in them in any permutation. (#c1,#c2)  
(6 marks)
- The number of punctuation symbols that are present in sentences that are greater than 10 words. (#d1,#d2)  
(2 marks)
- List of common occurrences of words in both the documents from part (a), (b), and (c) mentioned above. Example- list of common single-letter words in both documents, list of common 3-letter words in both documents and so on.  
(3marks)
- Draw a bar chart (shown below) to visually present this data. The data to be plotted is the set {(a,#a1,#a2),(b,#b1,#b2),(c,#c1,#c2),(d,#d1,#d2)}

(4 marks)

