

CP367: Assignment 3 – Winter 2024

Due on Feb 27, 2024 (Before 11:59 PM)

This is an **individual** assignment, and we will try to practice creating using regular expressions using C programming.

General Instructions:

- For this assignment, you must use C99 language syntax. Your code must compile using make **without errors**.
- **Test your program thoroughly with the GCC compiler in a Linux environment.**
- If your code does not compile, **then you will score zero**. Therefore, ensure you have removed all syntax errors from your code.
- **Gradescope** platform would be used to upload the assignments for grading. The link to the Gradescope assignment is available on Myls course page. Drag and drop your code file(s) for submission into Gradescope. **Make sure your file name is as suggested in the assignment; using a different name may score zero.**
- Please note that the submitted code will be checked for plagiarism. Submitting these .c files would confirm that you have not received unauthorized assistance in preparing the assignment. You also confirm that you are aware of course policies for submitted work.
- Marks will be deducted from any questions where these requirements are not met.
- Multiple attempts will be allowed, but your last submission will be graded before the deadline. Instructors reserve the right to take off points for not following directions.

Warning:

Follow the assignment instructions to the letter regarding the file names and function names, as this assignment will be auto-graded. If anything is not as described, the auto-grading fails, and your assignment will be given a **Zero** mark.

Question 1

Write a C program (name it as *regex.c*) that asks the user to enter the name of an email text input file (*emails.txt*) in the current directory. The program reads the text file and uses a regular expression with the `grep` command to redirect the valid emails to an output file (*valid.txt*). Use the `system()` function for running the shell command in the program.

The emails are considered valid:

- Usernames can contain letters (a-z), numbers (0-9), and periods (.).
- Username can contain: Periods (.), Hyphens (-), Underscores (_), Plus signs (+)
- Usernames cannot contain an ampersand (&), equals sign (=), apostrophe ('), comma (,), brackets (<,>), ! (exclamation mark), space, etc., or more than one period (.) in a row.

You are free to explore the error handling for file handling. Use the following command to compile the code:

```
$ gcc -Werror -Wall -g -std=gnu99 regex.c -o regex
```

The above code will create the `./regex` executable file

The expected output for executing:

When running on Linux Terminal, the command `./regex`. The following is the expected interaction with the program

```
● (r_40) ssehra@YVXV4FG67G Assignment_3 % ./regex
Enter the name of the input file: emails.txt
Enter the name of the output file for valid emails: valid.txt
Valid emails extracted and saved to valid.txt
⊗ (r_40) ssehra@YVXV4FG67G Assignment_3 % ./regex
Enter the name of the input file: invalid_input.txt
Error opening input file: No such file or directory
● (r_40) ssehra@YVXV4FG67G Assignment_3 % ./regex
Enter the name of the input file: emails.txt
Enter the name of the output file for valid emails: new_output.txt
Valid emails extracted and saved to new_output.txt
○ (r_40) ssehra@YVXV4FG67G Assignment_3 % █
```

Important Note: When submitting a source code file to Gradescope, make sure to name it like:

- `regex.c`