Dinesh Thapa   ID: 920879242
Github: dthapa770   CSC415 Operating Systems
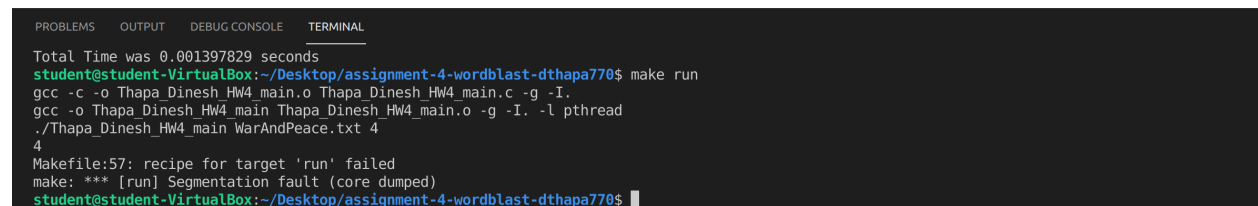
Assignment 4 – Word Blast

Description:

This assignment is to write a C program to read War and Peace and it is to count and tally each of the words that are 6 or more characters long using Linux file function i.e. open, close, read, lseek, pread.

Approach / What I Did:

For this assignment, I had to read a file name and number of threads which is given to us from the command line argument. I used array for my data structure that holds all the string after reading the file. Struct is defined to hold word and its occurrence, then the object is created from it that can hold array of words. I initialized my array which will allocate and store my data in my chosen structure. I divided my block size by given number of threads and started processing my threads. Each thread will call word process function. This is where important and critical process are computed. This function allocates buffer size and read the file descriptor into the buffer. String is broken down into tokens and they are processed. If the word length is greater than 5 and it matched with the word in array, it will enter the critical section where mutex lock is applied to avoid race condition and it will increment the number of count but if it doesn't match the string then again, we need to consider race condition and use mutex lock, we copy string into the array, and it will increment the count of the word. Then wait for the thread to finish, finally display the result in descending order. And as always, necessary cleaning up is done to clear any buffer, close file and destroy locks.
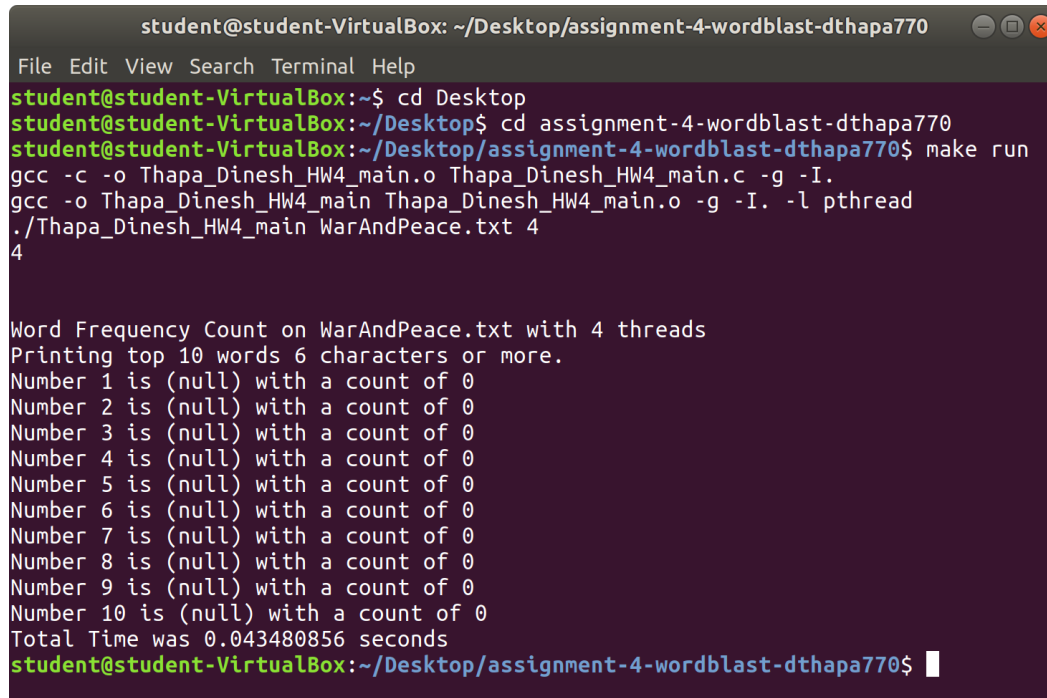
Issues and Resolutions:

Issue 1:


```
Total Time was 0.001397829 seconds
student@student-VirtualBox:~/Desktop/assignment-4-wordblast-dthapa770$ make run
gcc -c -o Thapa_Dinesh_HW4_main.o Thapa_Dinesh_HW4_main.c -g -I.
gcc -o Thapa_Dinesh_HW4_main Thapa_Dinesh_HW4_main.o -g -I. -l pthread
./Thapa_Dinesh_HW4_main WarAndPeace.txt 4
4
Makefile:57: recipe for target 'run' failed
make: *** [run] Segmentation fault (core dumped)
student@student-VirtualBox:~/Desktop/assignment-4-wordblast-dthapa770$
```

Resolution: Not sure what caused this segmentation fault. My guess is threads weren't handled correctly. I had to redo processing threads part and many changes.
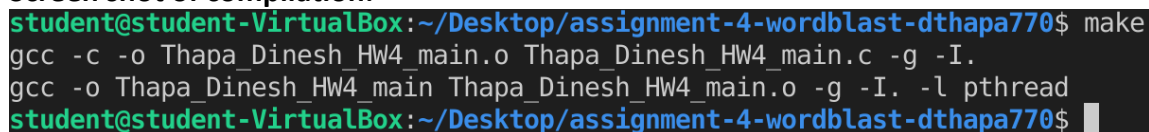
Issue 2:

```
                student@student-VirtualBox: ~/Desktop/assignment-4-wordblast-dthapa770

File  Edit  View  Search  Terminal  Help
student@student-VirtualBox:~$ cd Desktop
student@student-VirtualBox:~/Desktop$ cd assignment-4-wordblast-dthapa770
student@student-VirtualBox:~/Desktop/assignment-4-wordblast-dthapa770$ make run
gcc -c -o Thapa_Dinesh_HW4_main.o Thapa_Dinesh_HW4_main.c -g -I.
gcc -o Thapa_Dinesh_HW4_main Thapa_Dinesh_HW4_main.o -g -I. -l pthread
./Thapa_Dinesh_HW4_main WarAndPeace.txt 4
4


Word Frequency Count on WarAndPeace.txt with 4 threads
Printing top 10 words 6 characters or more.
Number 1 is (null) with a count of 0
Number 2 is (null) with a count of 0
Number 3 is (null) with a count of 0
Number 4 is (null) with a count of 0
Number 5 is (null) with a count of 0
Number 6 is (null) with a count of 0
Number 7 is (null) with a count of 0
Number 8 is (null) with a count of 0
Number 9 is (null) with a count of 0
Number 10 is (null) with a count of 0
Total Time was 0.043480856 seconds
student@student-VirtualBox:~/Desktop/assignment-4-wordblast-dthapa770$
```

Resolution: Solved with fixing word processing in my function. Count and words were not copied correctly.

Screen shot of compilation:

```
student@student-VirtualBox:~/Desktop/assignment-4-wordblast-dthapa770$ make
gcc -c -o Thapa_Dinesh_HW4_main.o Thapa_Dinesh_HW4_main.c -g -I.
gcc -o Thapa_Dinesh_HW4_main Thapa_Dinesh_HW4_main.o -g -I. -l pthread
student@student-VirtualBox:~/Desktop/assignment-4-wordblast-dthapa770$
```

Screen shot(s) of the execution of the program: