Honson Tran, Veton Abazovic
System Programming
Assignment #3
11/29/2017

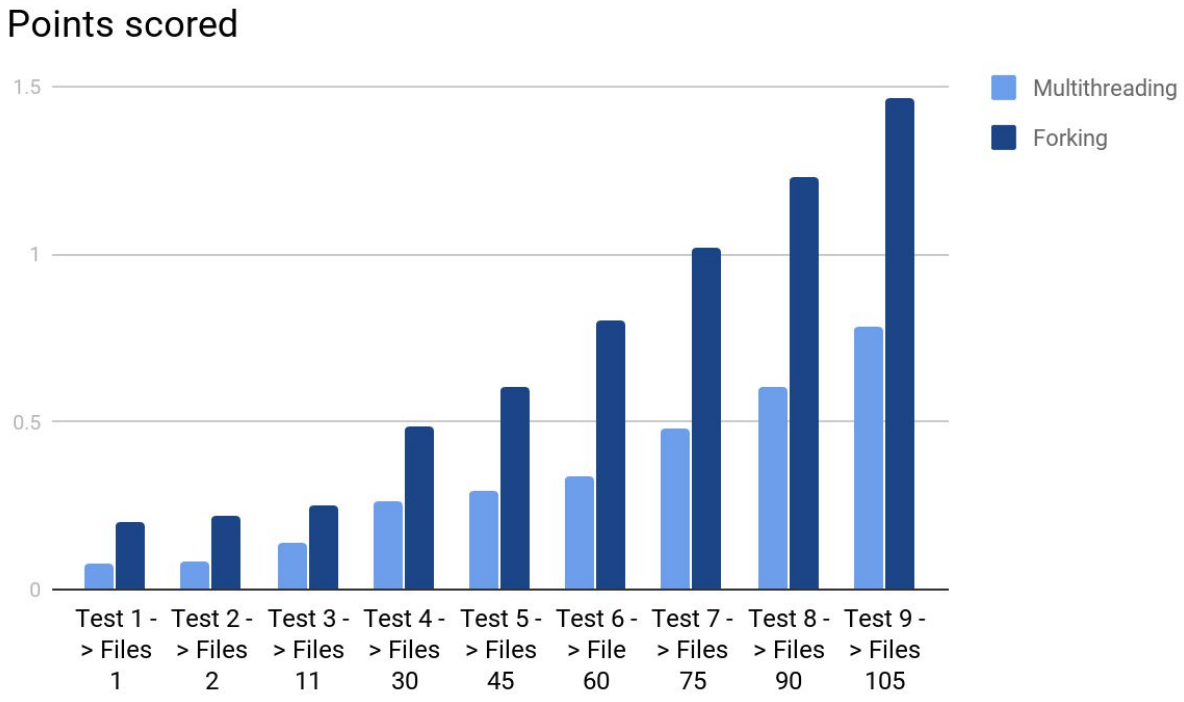
Basic Data Sorter - Multithreading

Assignment Details:

Project assignment 3 has us working on sorting multiple csv files in different directories by implementing multithreading instead of using multiple processors to compare how efficient one is over the other. When traversing through a directory the program will locate all the csv files in the directory and determine whether the csv file is valid (if the csv file contains the same header as movie_metadata.csv) then it will sort the file appropriately to the user's input. If a directory is found it will step inside and repeat the process. If the user doesn't specify an output directory then the code will sort the sorted csv file in the current directory. However, if the user does specify an output directory, the sorted files will be place there. If the output directory is not found during the time threading is over it will create the directory.

Findings:

For the assignment we continued to work with merge sort to sort all csv in our test directories. As we know the best/worst case for merge sort in term of time complexity is O(n log(n)), and in terms of space complexity worst case is O(n). After running a few test to compare the times of multithreading to multiple processors, the results are:



From the results above it is clear that multithreading is faster in comparison to having multiple processors. The machine that these test were run on was a ASUS ultrabook with the following specs: 512GB SDD, CORE I7, OS: Windows, Gtx 940, 16 GB DDR