

Andrew Hernandez & Michael Belmont

Project Part 1: A Multi-process Sorter

Design, Difficulties, and Assumptions

Coming into this project, we did not have much experience both navigating through the directories of the file system using c or multi-processing using fork. Experimenting with these two areas helped to initially write the basis for the code.

This program requires two parameters: `-c` (to specify sorting by columns) and a column name.

There are two optional flags: `-d` (followed by a specified directory to search through) and `-o` (followed by a specified directory to output to).

Assuming proper parameters, the program searches through the directory given (or the current directory if `-d` is not specified) and sorts each valid CSV within the directory and subdirectories.

Upon finding each subdirectory and each CSV, the program forks a child process to execute the desired operations. For each subdirectory, searching through it to find more CSVs. For each CSV, sort the CSV and output it either to the same folder if `-o` is not specified, or to the specified folder if `-o` is used.

How to use the program

Compile using: `gcc -g -Wall -Werror -fsanitize=address mergesort.c sorter.c -o sorter`

Then, run with: `./sorter -c columnName -d /folder/folder/folder -o /folder/folder/folder`