

CSC-213 Final Project: Sentiment Analysis

Leia, Haruko, Anil

Overview

- Program that analyzes the sentiment of txt files
- C for file access, concurrency, and processes
- Python for sentiment analysis
- Outputs a pie chart of emotions



Concepts

1. Files and File Systems
2. Processes
3. Parallelism

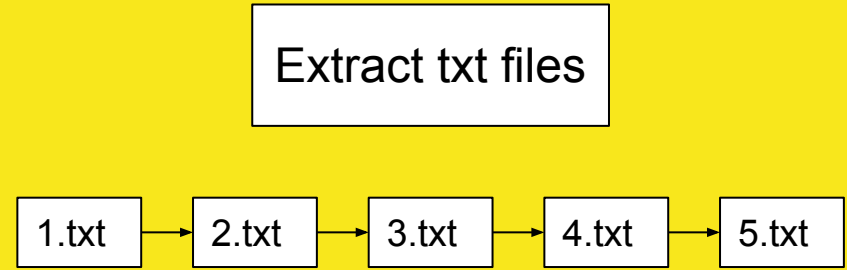
Files and File System

GOAL: Extract all .txt files from given directory

METHOD:

1. Open directory
2. Store readable txt files in a linked list
3. Recurse on directories

Libraries: <dirent.h>, <sys/stat.h>

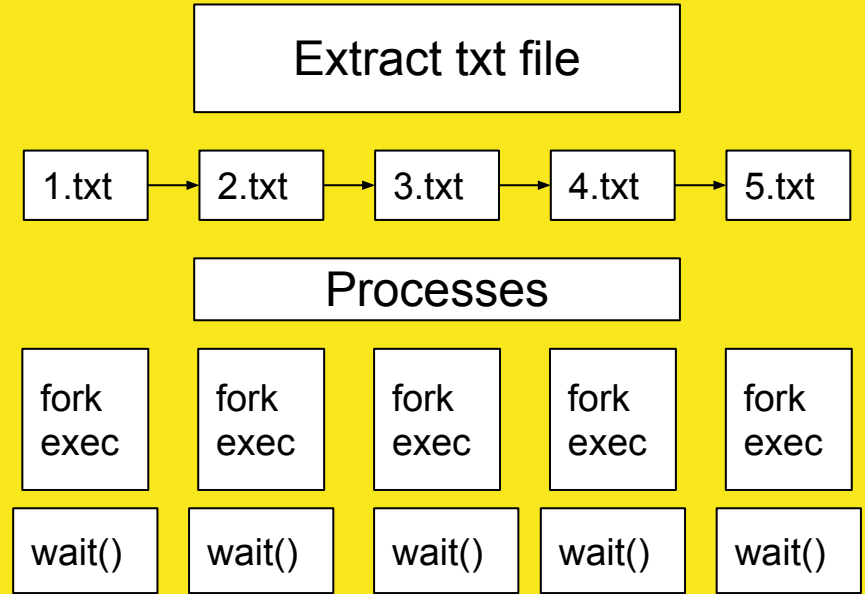


Parallelism Using Processes

GOAL: Run sentiment analysis concurrently

METHOD: Run processes concurrently through calling fork and exec in a loop. Call wait on each process in a different loop outside the first loop.

Libraries: <unistd.h>

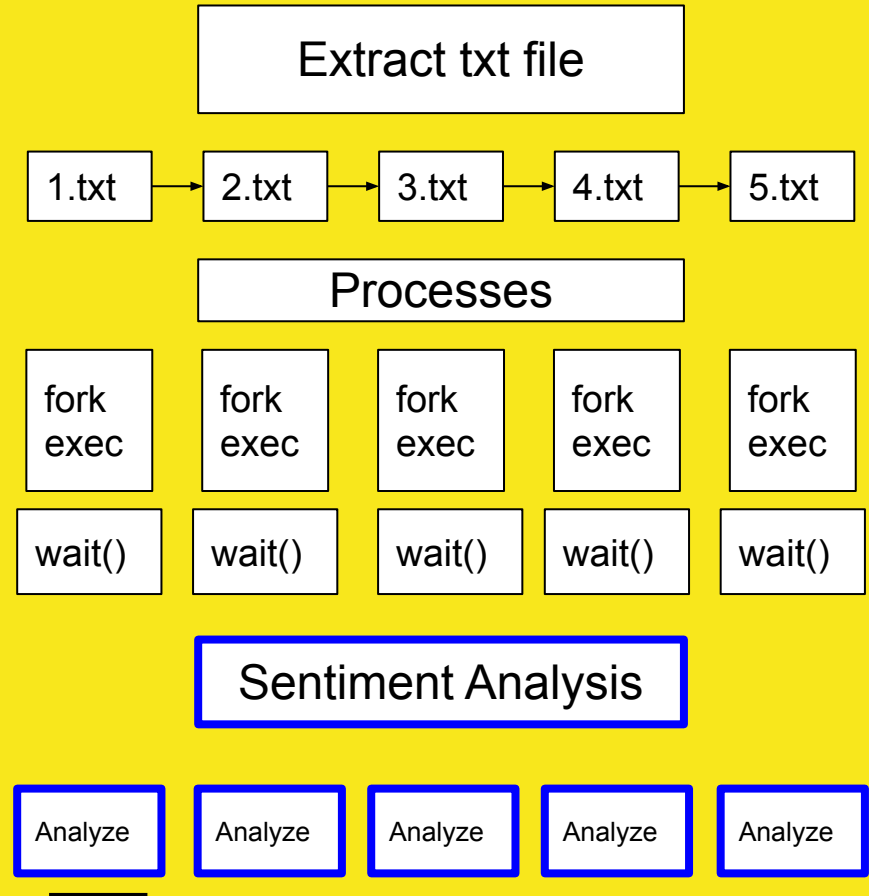


Sentiment Analysis

GOAL: Analyze and visualize the sentiment of the txt file.

METHOD: Reads the txt files and calls get_emotion method of text2emotion library that returns a Python dictionary labeling emotions (sad, anger, happiness, fear, surprise) to percentages for each file

Libraries: <text2emotion>,
<matplotlib>



Limitations and Future Works

- Program does not exist until all files are closed
- Limitation of text2emotion library
- Use pipe to rank the txt files in order of specific emotions

Thank you!
