

Word Statistics Processor Documentation

1. Overview:

There are 2 classes first is WordStatistics which is super class and WordStatisticsProcessor class .

2. WordStatisticsProcessor:

The WordStatisticsProcessor class is designed to process text files, either individually or within a directory, to gather statistics on word occurrences. It utilizes a multithreaded approach for efficient processing and updates a provided JTable with the computed word statistics.

3. Constructor:

WordStatisticsProcessor(JTable outputTable)

Constructor that initializes the processor with a reference to the JTable where the word statistics will be displayed. It also creates a pool of 5 threads.

4. Public Methods:

processFilesAsync(File fileOrDirectory)

Accepts a File parameter representing either a single file or a directory. Initiates the processing of files asynchronously using the provided ExecutorService. Updates the JTable with word statistics upon completion.

5. Private Methods:

-processFilesInDirectory(File directory)

Recursively processes all files in a given directory, submitting tasks to the executor service.

-processFile(File file)

Initiates the processing of a single file by creating a WordStatistics object and invoking the processTextFile method.

-processTextFile(File file)

Reads the content of a text file, updates a WordStatistics object, and returns the result.

-processLine(String line, WordStatistics wordStatistics)

Splits a line into words, updating the word statistics for word count, longest and shortest words, and specific word occurrences.

-displayWordStatistics(String fileName, WordStatistics wordStatistics)

updates the GUI (JTable) with the computed word statistics for a specific file.

5. WordStatistics Class (super class) :

Attributes:

wordCount: Total count of words in the file(s).

longestWord: The longest word encountered.

shortestWord: The shortest word encountered.

isCount, areCount, youCount: Counts for specific words ("is", "are", "you").

fileContent: The entire content of the processed file(s).

Methods:

getWordCount(), getLongestWord(), getShortestWord(): Accessor methods for statistics.

getIsCount(), getAreCount(), getYouCount(): Accessor methods for specific word counts.

getFileContent(), setFileContent(String fileContent): Accessor and mutator for file content.

incrementWordCount(int count): Increments the total word count.

updateLongestShortestWord(String word): Updates the longest and shortest word based on input.

updateWordOccurrences(String word): Updates specific word occurrences.

Team members :

- | | |
|---------------------------------|-----------|
| 1- Shams Ahmed Shaalan | 202000439 |
| 2- Maria Ashraf ashy | 202000695 |
| 3- Mahmoud Ahmed Mohamed Moomen | 202000837 |
| 4- Yehia Mohamed Ahmed Hassan | 202001062 |
| 5- Mohamed Ebrahim Abdelhameed | 202000717 |
| 6- Omar Hamdy Orany | 202000591 |