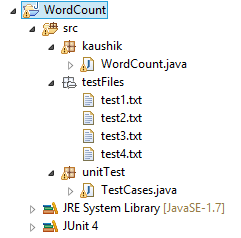
Word Counter

Below is the folder structure for the word counter problem. Code is written using Eclipse IDE.



Supporting files and a zipped file is provided in same directory of readme.

This file structure or zipped folder can directly be imported in ide to make it work.

Project contains –

WordCount.java – main class to count words.

Test files for tested cases

Testcases.java – Junit class to test all four test cases.

I have tested the code with four distinguished inputs with plain text, numeric values and as well as UTF-8 characters.

This version of code will not count UTF-8 characters.

The testing done for code is sufficient due to the fact of code’s full proof design.

Code is developed to accept the characters which user wants to count for.

I have created a set of characters to monitor to make a valid word and hence count it. This approach is little bit slower than traditional way but it has room of improvement.

As the code process each line from the file at a single time. This approach can ultimately be scaled be parallelism. Consider situation where there are 40 lines to read from a line.

I can start 4 different threads to start the same logic of computation and update the counter.

The counter is publically accessible and each update from thread will be incremental update therefore the output will be thread safe and reliable.

WordCount.java contains function “calculateWords(String file)” it is has a public it call and can be assessed throughout the project.