**Lab 5\_3**

**SpellCheck Application**

For the SpellCheck application, various Collection classes can be used to store the dictionary of words – the words are read in from ‘words.txt’. It counts the number of misspelt words found in the text you are spell checking (alice30.txt ). A larger text file war-and-peace.txt is also given.

Use IntelliJ Profiler to generate % of time and actual time (in ms) for contains() method of your chosen Collection class – code as given uses a LinkedList.

Complete the following table:

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Structure used to store dictionary** | **% of Time for contains() method** | **Time in ms for contains() method** | **Big Oh for contains() method**  **n – size of the dictionary -10000** |
| LinkedList | Higher than others | 83 ms | O(n) |
| ArrayList | Lower than LinkedList, higher than sets | |  | | --- | |  |   43 ms | |  | | --- | | O(n) | |
| HashSet | Very minimal | 1 ms | O(1) average, O(n) worst |
| TreeSet | Minimal | 0 ms | O(log n) |
|  |  |  |  |

Obtained with Intel(R) Core(TM) i7-10870H CPU @ 2.20GHz 2.21 GHz, Java Version \_\_\_, Windows 10

Try different Collection classes and see the different values you will get for the contains() method. Use the larger text file if the smaller file give values that are too small.

What Collection class would you recommend for the SpellCheck application?

\_My recommendation is HashSet.\_\_

Explain your answer \_\_\_I prefer to use HashSet because it is really fast. Of course it is slower than TreeSet but it is really good with larger files. It uses less memory than TreeSet and, this is also good thing. It has really good and consistent performance. When we look all of the facts HashSet is better than others.\_\_\_