

Question 1: Removing stop-words from a given text file

For the CM1210 2nd Java coursework, the first question had us make a function to remove a set amount of stop-words from a text file given to us by the lecturer – this function can be found as “removeStopwords(GPT3, stopwords)” within Main.java. Firstly, I made sure that the two files given to us in this assignment could be inputted into the function, “GPT3” is where the user would specify (by hardcoding in) the file location for the GPT3 text file, and “stopwords” is where the user would specify (by hardcoding in) the file location for the stopwords text file. I also had the function “throw” a FileNotFoundException error this was needed in case the function could not find the text files provided within the code. If I were to implement this within a professional setting, the user would be able to provide the location of the file(s) through the command line, instead of accessing the Java file and entering it through there.

The question was quite open-ended concerning how much editing could be made to the GPT3 text file (apart from removing the stopwords provided). Within my function, I had used regex to remove all punctuation, keep all numbers & convert all words to lower-case. I did this because I noticed that all the strings/words within the stopwords text file were lower-case, so I made sure that the text within the GPT3 text file was standardized to lower case. When outputting the ArrayList from the “removesStopwords()” procedure, this is what you get:

[illegible]

The second question given to us in this coursework was to implement the Insertion Sort algorithm into Java. The pseudocode provided within the lectures was more than sufficient to allow me to translate it into code within Java. I made sure that the procedure had only

one input "GPT3" which must be an ArrayList. In the screenshot below, this is what the output of my InsertionSort() procedure looks like when using the ArrayList outputted in Question 1 :

As mentioned in the Question 1 section above, the numbers seen within this output are sorted alphanumerically – this means that the numbers are sorted by using the 1st digit, while the words are sorted alphabetically.

The third question given to us in this coursework had us implement the Merge Sort algorithm into Java. Personally, I found the pseudocode given to us from the lectures a little lacking when trying to translate it directly into Java, however, the lecture slides given to us was very helpful in allowing me to implement recursion into my procedure. My main concern with this question was to make sure the `mergeSort()` procedure has only one input (the GPT3 ArrayList from Question 1). The examples shown within the lecture had procedures that needed 2 additional parameters. However, I was able to overcome this difficulty by implementing a “helper function” called “merge” that would be used to sort the multiple lists created by the main `mergeSort()` function. Learning from the lectures, I made sure to implement recursion when splitting the inputted ArrayList “GPT3” within `mergeSort()`. The output for the function can be seen below:

Like in the output for the previous question, the numbers are sorted alphanumerically, and words are sorted alphabetically, both insertSort & mergeSort have the same output as intended.

For the fourth question, we were given the task to create a “performance” function. This function would essentially test the speed (time taken) & the number of swaps/moves that both insertionSort & mergeSort would have every one hundred (e.g. 100, 200, 300 until the max array size). To do this I decided to create a “performanceList” ArrayList, and within a loop, add each string/item from the GPT3 list outputted by Question 1 into the performanceList and increase a counter by 1 for each occurrence of the loop. I would then

check if the counter is able to be divided by a hundred (to check if there are 100/200/300 etc) items within the list and then I would run both sorts. I have also implemented when the counter equals the GPT3 ArrayList size it will perform the same performance check of speed and swaps/moves. To measure the speed, I used Java 8's inbuilt "System.nanoTime()" to measure the time difference between before and after the specific sort procedure is called. I then found the difference between these two numbers and outputted them as nanoseconds as that is the default output of using nanoTime. To count the swaps/moves of the two different sorting functions, I added insertion swap and merge move counters into the sorting functions. Anytime that a swap or move occurred, the counter variable would increment by one. The output for this function can be seen below:

```
>>>Sort Size: 1200
Insertion Swap Duration = 1213301 nanoseconds
Insertion Swap No. of swaps performed = 366196
Merge Swap Duration = 1975300 nanoseconds
Merge Swap No. of moves performed = 12352

>>>Sort Size: 1300
Insertion Swap Duration = 1247599 nanoseconds
Insertion Swap No. of swaps performed = 427781
Merge Swap Duration = 2946300 nanoseconds
Merge Swap No. of moves performed = 13552

>>>Sort Size: 1400
Insertion Swap Duration = 1277700 nanoseconds
Insertion Swap No. of swaps performed = 491212
Merge Swap Duration = 2401099 nanoseconds
Merge Swap No. of moves performed = 14752

>>>Sort Size: 1500
Insertion Swap Duration = 1423899 nanoseconds
Insertion Swap No. of swaps performed = 559679
Merge Swap Duration = 2306399 nanoseconds
Merge Swap No. of moves performed = 15952

>>>Sort Size: 1600
Insertion Swap Duration = 1549200 nanoseconds
Insertion Swap No. of swaps performed = 633192
Merge Swap Duration = 2379300 nanoseconds
Merge Swap No. of moves performed = 17152

>>>Sort Size: 1700
Insertion Swap Duration = 1584400 nanoseconds
Insertion Swap No. of swaps performed = 718298
Merge Swap Duration = 2674900 nanoseconds
Merge Swap No. of moves performed = 18352

>>>Sort Size: 1800
Insertion Swap Duration = 2863300 nanoseconds
Insertion Swap No. of swaps performed = 809313
Merge Swap Duration = 3000500 nanoseconds
Merge Swap No. of moves performed = 19552

>>>Sort Size: 1900
Insertion Swap Duration = 1752299 nanoseconds
Insertion Swap No. of swaps performed = 900676
Merge Swap Duration = 3268300 nanoseconds
Merge Swap No. of moves performed = 20752

>>>Sort Size: 2000
Insertion Swap Duration = 1826300 nanoseconds
Insertion Swap No. of swaps performed = 1000703
Merge Swap Duration = 3285700 nanoseconds
Merge Swap No. of moves performed = 21952

>>>Sort Size: 2100
Insertion Swap Duration = 1920999 nanoseconds
Insertion Swap No. of swaps performed = 1105031
Merge Swap Duration = 3268600 nanoseconds
Merge Swap No. of moves performed = 23204

>>>Sort Size: 2200
Insertion Swap Duration = 2040600 nanoseconds
Insertion Swap No. of swaps performed = 1207689
Merge Swap Duration = 3425099 nanoseconds
Merge Swap No. of moves performed = 24504

>>>Sort Size: 2236 [MAX SIZE]
Insertion Swap Duration = 687700 nanoseconds
Insertion Swap No. of swaps performed = 1243148
Merge Swap Duration = 3422101 nanoseconds
Merge Swap No. of moves performed = 24972
```

The results seen here corollate with the space-time complexity of both insertion & merge sort. With the max size being 2236 elements, the time difference taken between the two is very large, this is because merge sort has a complexity of $O(n \log(n))$ while insertion sort has a complexity of $O(n^2)$ – which also explains the large difference between the swaps taken by insertion sort and moves taken by merge sort.

Question 5: Creating a Circular Array Queue

For the last question within our assignment, we were given skeleton code and had to fix it to generate a working Circular Array Queue. The first problem I ran into was that I was unsure of how to handle if the queue had hit the max size limit. After reviewing the lectures, I found out that I would have to create a new queue double the size of the old queue, copy the elements of the old queue into the newer, bigger queue, and then set the new queue as the default/primary queue. The lecture slides had instructed us to use “System.arraycopy()” to move the contents of the old queue into the new one. However, I was unable to implement this properly as I could not get it to work, so instead, within the “enqueue()” function, I decided to use a loop to add each element from the old queue into the new queue (which was double the size plus 1 of the old queue), and then set the new queue as the default and reset the front and rear pointers accordingly. To help me achieve this I had to create an additional function called “isFull()”. This function would return true if the queue was full, and false otherwise. Within the loop, it would check if the queue was full, and if it was, it would append each item from the old queue into the new queue by dequeuing each item which can be seen in the screenshot below:

```
public void enqueue(Object theElement)
{
    if (isFull()) {
        Object [] queueNew = new Object[(queue.length * 2)+1]; //if queue is full, create a new queue with double the size + 1
        int indexNew = 1;
        while (!isEmpty()) {
            queueNew[indexNew] = dequeue();
            indexNew++; //while isEmpty == False, append (by de-queueing) each item from the old queue to the new queue and increment indexNew by 1
        }
        queue = queueNew;
        front = 0;
        rear = indexNew - 1; //re-set the new queue with larger storage space as the new queue and reset front pointer to 0 and rear to the indexNew -1 (since arrays are zero-based)
    }
    rear = (rear + 1) % queue.length;
    queue[rear] = theElement;
} //else if the queue is not full, set the rear to the true modulus of rear & queue.length (rear + 1 because Java treats the modulus operator differently) & then set the rear of the queue to the inputted element
```

For the dequeue function, I had followed the lecture notes/pseudocode given and I did not deviate from it. The correct output of my Circular Array Queue Java code can be seen below:

```
Rear element : element12
Front element : element3
Removed element: element3

Rear element : element12
Front element : element4
Removed element: element4

Rear element : element12
Front element : element5
Removed element: element5

Rear element : element12
Front element : element6
Removed element: element6

Rear element : element12
Front element : element7
Removed element: element7

Rear element : element12
Front element : element8
Removed element: element8

Rear element : element12
Front element : element9
Removed element: element9

Rear element : element12
Front element : element10
Removed element: element10

Rear element : element12
Front element : element11
Removed element: element11

Rear element : element12
Front element : element12
Removed element: element12

empty queue
```