

PRACTICAL ASSIGNMENT

PROBLEM SOLVING WITH DATA STRUCTURES AND ALGORITHMS

PERFORMANCE ANALYSIS ON SORTING ALGORITHMS

Assignment Description

Sorting is considered one of the most frequently performed operations on data. Sorting elements i.e. integer, float, or string in a dataset (especially a very large dataset) can be very time consuming if inefficient algorithm or data structure is applied. In general, sorting a set of elements can be executed using the comparison-based or non-comparison-based algorithms. When a comparison-based algorithm is used, elements in an input data structure are compared in pairs and then elements will be swapping around based on the comparison results. On the other hand, non-comparison-based algorithm does not compare elements directly and then put them in order. Instead, it applies other techniques to achieve sorting. In this assignment, every group is required to perform a thorough comparative analysis between these two categories of sorting algorithms.

Summary of Assignment Tasks

- Apply at least **FOUR (4) sorting algorithms** (must be **3 comparison-based** and **1 non-comparison-based** methods) on datasets with **different input sizes (n)**. You can start with $n = 100k$, and gradually increases the size up to $100M$.
- Data that used for sorting are **numbers** and **English words**.
- Data will be placed in two data structures, i.e. **array list** and **linked list**
- Every student group will be assigned with **FOUR (4) sorting algorithms (from the following table)** by the lecturer, and you may view the assigned algorithms in the shared excel file.

No.	Comparison-based	No.	Non-comparison-based
C1	Selection-sort	N1	Bucket-sort
C2	Insertion-sort	N2	Radix-sort
C3	Bubble-sort	N3	Counting-sort
C4	Shell-sort		
C5	Heap-sort		
C6	Merge-sort		
C7	Quick-sort		
C8	Tim-sort		
C9	Comb-sort		

The followings are the key tasks your group has to implement and complete in the assignment.

Task 1 (10%):

Generate TWO (2) large datasets, one with at least 10 million numbers, and another one with at least 1 million English words. The data could be generated by defining a method/ function or download from the online resources (please specify the resource clearly in the report).

Task 2 (20%):

Create a Java project and then write Java program(s) to implement the FOUR (4) sorting algorithms with TWO (2) data structures.

	Number Data		English Word Data
Sorting Algorithm	Array List	Linked List	Array List
Comparison-based algo1	√	√	√
Comparison-based algo2	√	√	√
Comparison-based algo3	√	√	√
Non-comparison algo1	√	√	√

Task 3 (5%):

Ensure the correctness of the implementation. Output the results to the lecturer by utilizing small example size. This is for the lecturer to inspect whether or not the implemented algorithms and data structures are working accurately.

Task 4 (15%):

Perform numerous experimental tests on sorting (with different element sizes, randomness, with/without duplicates, etc.) using the four algorithms, and get the elapsed times (average case & worst case) for the algorithms.

Task 5 (40%):

Prepare a report (group-based) to include the experiment results that can be used to perform a comparative analysis for the algorithms and data structures. Conclude your findings in the report. The report's length shall cap at 15 pages.

Task 6 (10%):

Presentation to showcase your work. Every student in the group must take part in the presentation.

Further Information

Prepare the necessary files (data and source files) for creating the data structure(s) and algorithms needed in the assignment. You may give additional assumptions for your assignment. To make your program more robust and avoid problems at run time, do as much status/error checking as you could in your program. And, good organization of the code and meaningful variable names would help readability, and liberal use of comments can help the marker understands what the program does and why.

Assessment and Submission

This is a group assignment. Form a group of 4 members, preferably from the same programme as yours (not compulsory). You may register your group in the google sheet provided in the link below:

https://utarict.sharepoint.com/:x:/r/sites/UECS2083_U ECS2413_U ECS2453_Jun2024/Shared%20Documents/General/Assignment_Grouping_U ECS2083_U ECS2413_U ECS2453_2406.xlsx?d=w7b3c25b93d3f4b9c89ac2d260e15db62&csf=1&web=1&e=gDQYYV

Deliverables:

(1) Java Project Folder

Should contain both *src*(.java files) and *bin*(.class files) folders.

If there's any jar file, indicate the name of the jar file and link to download the jar file.

Assessment Criteria (Refer to 'Assignment marking sheet'):

- Framework Design (Use of interfaces and abstract classes)
- Classes for storing objects (data structures/containers)
- Search Algorithms (should have 4 algorithms)
- Test program (main program, set of evaluations using the algorithms and data structures)
- Exception and error handling (if any)
- Presentation of source codes

(2) Report

Attach the "Assignment marking sheet" as the FIRST PAGE of your report.

The **REPORT SHOULD CONTAIN** the following:

- Short introduction
- Proposed solution – data structures and algorithms used for the search problem, discuss the time complexities (Big-O) of the algorithms
- Sample of input data and test cases
- Sample output(s) or results of your program – *Tables, Screenshots, etc.*
- Short discussion on the results
- Conclusion

(3) Presentation

- Physical or online presentation
- Each member must present
 - Member #1: Introduce yourself with your camera is on (name, student ID, programme), Followed by presentation
 - Member #2: Introduce yourself with your camera is on (name, student ID, programme), Followed by presentation
 - Member #3: Introduce yourself with your camera is on (name, student ID, programme), Followed by presentation
 - Member #4: Introduce yourself with your camera is on (name, student ID, programme), Followed by presentation
 -
- Duration: ~20 mins
- Content of presentation:
 1. Short introduction and explain about your datasets
 2. Explain if you have any assumption or additional criteria
 - 3.1 Explain sorting algorithm 1

- 3.2 Explain your design & data structure used
- 4.1 Explain sorting algorithm 2
- 4.2 Explain your design & data structure used
- 5.1 Explain sorting algorithm 3
- 5.2 Explain your design & data structure used
- 6.1 Explain sorting algorithm 4
- 6.2 Explain your design & data structure used
- 7. System Demo (show results)
- 8. Conclusion

Submission:

Due date: **W12: 06/09/2024 (Friday)**

1. Java Project Folder and Report

Create a folder with group number and copy both (1) and (2) into the folder. Ex: 'Group Number 1'. Zip the folder and submit it through link created in WBLE.

2. Presentation

- File to submit: ppt slide
- Submit it through the link created in WBLE
- Subject: <CourseCode_Assignment Presentation_GroupNumber> ex: UECS2453_Assignment Presentation_Group1

Total mark

The total mark of this practical assessment is 100. The 100 marks will contribute 20% of your final mark. The report will be marked for *correctness*, *completeness*, *presentation style*, and *relevant use of diagrams/tables/graphs*, etc. And the program implementation will be marked for *correctness*, *completeness*, *adequate testing*, and *documentation/comments*. It's your responsibility to understand the requirements of the tasks and prepare well for your submission. You might be asked questions about the works you submit to ensure that you understand them.

Late Submission

No late submission of assignment is allowed. Assignment received after the due date without valid reasons will be penalized using the following policy: 5 marks will be deducted for every day the assignment is overdue.

Plagiarism

It is important that your solutions to the practical assignment be your own work. It is perfectly acceptable to seek help and advice when completing the practical assignment, but this must not be taken to the point where what is submitted is in part someone else's work. Any group found with high plagiarism rate in the submitted work will be graded as fail in the assignment.