
CS246: Database Management Systems Lab

Lab # 02 (1 Questions, 100 Marks)

Lab session: AL1

Held on: 16-Jan-2023 (Mon)

Lab Timings: 14:00 to 17:00 Hours Pages: 3

Submission time: 18:00 Hrs, 16-Jan-2023

Instructors Dr. V. Vijaya Saradhi & Prof. Jatindra Kumar Deka

Head TAs Adithya K Moorthy and Laxita Agrawal

Department of CSE, IIT Guwahati

Question 1: (100 points)

Implement in C Programming Language the following problem. Use of any programming language other than C is not allowed and will lead to awarding 0 marks.

Problem Statement Given input data files containing one data item per line, your task is:

Task 01 - Input generator Write input generator program having three distinct functions to generate several sets of input files as per the following specification:

Task - 01 (a) Integer input generator - 4 Marks Randomly generate

1. 10,000 integers between 0 and 1000000. Each line containing one integer.
2. Save them in a file i10k.txt
3. Repeat steps (1) and (2) by increasing the input size by a factor of 2. That is 20,000 integers; 40,000 integers; 80,000 integers etc. Save the output into files i20k.txt, i40k.txt, i80k.txt etc.
4. Generate 10 such inputs.

Task 01 (b) - Date input generator - 10 Marks Randomly generate

1. 10,000 dates between 15-Aug-1947 and 01-Jan-2023. Each line containing one date. The date format is: dd-mmm-yyyy.
2. Save them in a file d10k.txt
3. Repeat steps (1) and (2) by increasing the input size by a factor of 2. That is 20,000 dates; 40,000 dates; 80,000 dates etc. Save the output into files d20k.txt, d40k.txt, d80k.txt etc.
4. Generate 10 such inputs.

Task 01 (c) - Date and time stamp input generator - 10 Marks Randomly generate

1. 10,000 date and time stamps between 15-Aug-1947 00:00:00 and 09-Jan-2023 23:59:00. Each line containing one date and time. The date and time format is: dd-mmm-yyyy hh:mm:ss
2. Save them in a file dt10k.txt
3. Repeat steps (1) and (2) by increasing the input size by a factor of 2. That is 20,000 date and time items; 40,000 date and time items; 80,000

date and time items etc. Save the output into files dt20k.txt, dt40k.txt, dt80k.txt etc.

4. Generate 10 such inputs.

Task 02 - Read each

1. Task (2a) (2 Marks) integer file (i10k.txt, i20k.txt, etc)
2. Task (2b) (2 Marks) date file (d10k.txt, d20k.txt, etc)
3. Task (2c) (2 Marks) date and time file (dt10k.txt, dt20k.txt, etc)

Task 03 - Sort the data using (a) Quick sort (b) Bubble sort. That is

1. Task (3a) (5 Marks) Sort 10 integer input files using Quick sort
2. Task (3b) (5 Marks) Sort 10 integer input files using Bubble sort
3. Task (3c) (10 Marks) Sort 10 date input files using Quick sort
4. Task (3d) (10 Marks) Sort 10 date input files using Bubble sort
5. Task (3e) (10 Marks) Sort 10 date and time input files using Quick sort
6. Task (3f) (10 Marks) Sort 10 date and time input files using Bubble sort

Task 04 - Write the sorted data to a new output file. Names the output files as described:

Task 04 (a) (5 Marks) Quick sort

1. qi10k.txt when input file contains 10000 integers; qi20k.txt when input file contains 20000 integers; etc.
2. qd10k.txt when input file contains 10000 dates; qd20k.txt when input file contains 20000 dates; etc.
3. qdt10k.txt when input file contains 10000 date and time stamps; qdt20k.txt when input file contains 20000 date and time stamps; etc.

Task 04 (b) (5 Marks) Bubble sort

1. bi10k.txt when input file contains 10000 integers; bi20k.txt when input file contains 20000 integers; etc.
2. bd10k.txt when input file contains 10000 dates; bd20k.txt when input file contains 20000 dates; etc.
3. bdt10k.txt when input file contains 10000 date and time stamps; bdt20k.txt when input file contains 20000 date and time stamps; etc.

Task 05 - Measure the time taken for

Task 05 (a) (3) sorting alone

Task 05 (b) (3) sorting and writing to output files.

Task 06 - Output (4) Your program output should be of the following format.

The following table should be written to an output file `roll_number-output.txt`

Quick Sort		
Description	Only sorting	Sorting and writing into file
Integers 10K	0.100 sec	0.200 sec
Integers 20K	0.200 sec	0.400 sec
...
Date 10K	0.200 sec	0.400 sec
Date 20K	0.400 sec	0.800 sec
...
Date & time 10K	0.200 sec	0.400 sec
Date & time 20K	0.400 sec	0.800 sec
Bubble Sort		
Integers 10K	0.100 sec	0.200 sec
Integers 20K	0.200 sec	0.400 sec
...
Date 10K	0.200 sec	0.400 sec
Date 20K	0.400 sec	0.800 sec
...
Date & time 10K	0.200 sec	0.400 sec
Date & time 20K	0.400 sec	0.800 sec

Instructions Adhere to the following

File naming Prepend C program file names with your roll number. Adhere to the input and output file naming convention as given in the problem description.

Independent efforts You should make an honest and independent effort in obtaining the solution to the above problem. You are also encouraged to bring one data structures and algorithms text book and one C programming text book of your choice.

Discussions with fellow students are not allowed.

Internet Use of internet during lab hours is not allowed.

Evaluation At the end of 18:00 hours, TAs will come and evaluate your program. Leave the lab once your evaluation is completed.

Marking Scheme The evaluation criteria is as follows:

Task 01 24 marks for input generator

Task 02 6 marks for reading input data into appropriate data structures.

Task 03 50 marks for sorting the input data.

Task 04 10 marks for writing the output of sorting into output files

Task 05 6 marks for measuring the time taken

Task 06 4 marks for obtaining the measurement table