Birla Institute of Technology & Science, Pilani
2nd Semester 2016-17 - CS F211 – Data Structures and Algorithms

Lab 3 – 31st Jan 2017
Topics –Exploring maximum size of dynamic array, Sorting

1. Measuring the maximum size of dynamically allocated array

In this exercise you need to find the maximum “allocatable” size of a one dimensional array of type integer on your system. You can iteratively allocate (and free) a dynamic array while doubling the size every time until you incur failure.

2. Sorting very large files

Given a very large file of few GBs of data (few billions of records stored in it). You need to sort this file. We came to know from the above exercise, the maximum size you can allocate to an array. So obviously you can’t make an array of size of a few billions. However, you can merge two sorted files of smaller size and make a larger sorted file using FILE I/O.

Use this technique to design a sorting algorithm to sort very large files: Use merge sort at the top level to create multiple files of smaller sizes. Sort these smaller files using insertion sort and then merge all of them (two at a time) in a hierarchical manner to get the bigger sorted file. You can assume that the number of records in the file is a multiple of 2.

After sorting, call a function `verify(P, N)` which prints all N sorted numbers starting from Pth index (Indexing start with 1, for printing).

Dataset generation script:

- Create a file `data_generator` and write following code –

```
for i in $(seq $1)
do
     echo $RANDOM
done
```

- Set its permission to be executable

```
chmod +x data_generator
```

- Execute it by passing required data size as argument (Be careful with the file size)

```
./data_generator 10000000 > input.txt
```

3. Profiling

Measure the time taken and space used by above problem by using the techniques referred to in the last lab session.