

DESIGN AND ANALYSIS OF ALGORITHMS

[Bcse 2350]

ANKIT VERMA
Batch II

MODULE I

INTRODUCTION TO DAA

Algorithm

Algorithm is a step by step method to solve a problem. It refers to a set of rules/instructions that step by step define how a work is to be executed upon in order to get the expected results.

Characteristics of Algorithm

- (*) Input :- Algorithm must contain '0' or more input.
- (*) Output :- Algorithm must contain '1' or more output.
- (*) Finiteness :- Algorithm must complete after finite no. of steps.
- (*) Definiteness :- The step of the algorithm must be defined precisely or clearly (unambiguous).
- (*) Effectiveness :- The step of an algorithm must be effective (can be done successfully).

Analysis of Algorithm

Analysis of algorithm depend upon the time & memory taken by the algorithm.

- i.e
- (*) Time Complexity
 - (*) Space Complexity

Time Complexity

The amount of time required by the algorithm is called time complexity.

Space Complexity

The amount of space / memory required by the algorithm is called as Space Complexity.

Application of Algorithm

- (*) Bioinformatics (*) Finding shortest root in road map
- (*) Internet
- (*) Ecommerce (*) Product of Matrices.

Growth functions

	2^n	$n!$
$n=1$	2	1
$n=2$	4	2
$n=3$	8	6
$n=4$	16	24
$n=5$	32	120

$n! > 2^n \text{ for } n > 4$

$n!$ has higher growth rate than 2^n

Arrange the following according to increasing growth rate order.

$n, 2n, n\log n, n!, n^3, n^5, n^2, 1$

\Rightarrow constant \leq logarithm \leq polynomial \leq Exponential \leq factorial

Ans: $1, n, n\log n, n^2, n^3, n^5, 2^n, n!$

Asymptotic Notation

These notations are used to describe the asymptotic running time (when the input size n is very large i.e $n \rightarrow \infty$) of an algorithm which defines functions whose domains are the set of Natural Numbers.

⇒ The time complexity of a function can be represented by asymptotic notations.

⇒ The different asymptotic notations are :-

- (*) Big Oh (O)
- (*) Big omega (Ω)
- (*) Big theta (Θ)
- (*) Little Oh (\mathcal{O})
- (*) Little omega (ω)

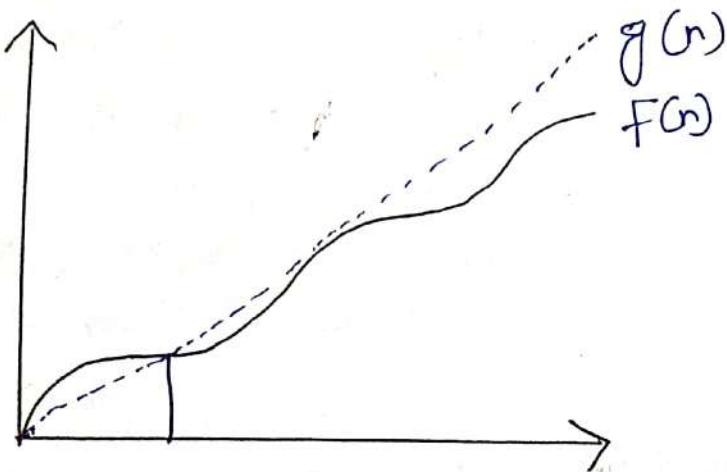
Big Oh Notation, O

The Notation $O(n)$ is the formal way to Express the upper bound of an algorithm's running time. It measures the worst case time Complexity on the longest amount of time an algorithm can possibly take to complete. A function $f(n)$ can be represented is the Order of $g(n)$ that is $O(g(n))$, if there exists a value of positive integers n as n_0 and a positive constant C such that

$$f(n) \leq C \cdot g(n) \quad \text{for } n > n_0 \text{ in all cases}$$

$g(n) \rightarrow$ upper bound for $f(n)$

$$f(n) \in O(g(n))$$



$$f(n) \leq c \cdot g(n)$$

Example:- $f(n) = 4 \cdot n^3 + 10 \cdot n^2 + 5 \cdot n + 1$

let $g(n) = n^3$ $\{c=4\}$

$f(n) \leq 5 \cdot g(n)$ for all values of $n > 2$

Complexity of $f(n) \Rightarrow O(g(n))$
i.e $O(n^3)$

Big Omega (Ω)

The Notation $\Omega(n)$ is the Normal way to Express the lower bound of an algorithm's running time. It Measures the best case time complexity or the best amount of time an algorithm can possibly take to Complete.

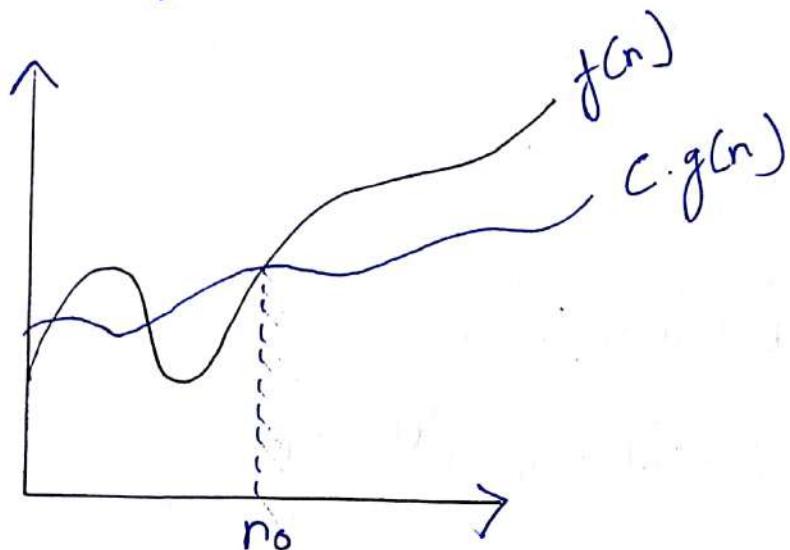
$$f(n) = \Omega(g(n)) \text{ when there exists constant}$$

i.e $f(n) \geq c \cdot g(n)$ for larger value of n .

where g is a lower bound for fun f

$$f(n) \geq c \cdot g(n)$$

$$f(n) \in \Omega(g(n))$$



Example:- $f(n) = 4 \cdot n^3 + 10 \cdot n^2 + 5 \cdot n + 1$

$$g(n) = n^3$$

$$f(n) \geq 4 \cdot g(n) \quad \forall n > 0$$

Complexity of $f(n) = \Omega(g(n))$
i.e. $\Omega(n^3)$

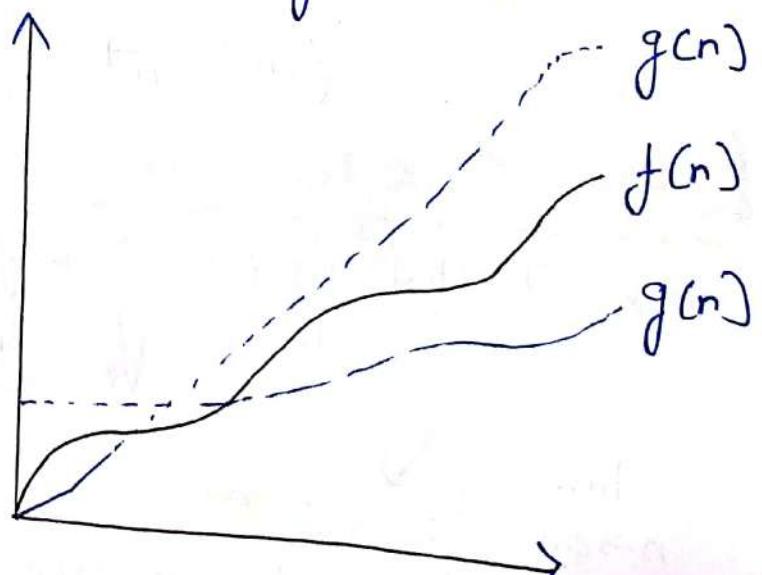
Big Theta (Θ)

The Notation $\Theta(n)$ is the formal way to Express both the lower bound and the upper bound of an algorithm's running time.

$$f(n) = \Theta(g(n))$$

when there Exist constant c_1 & c_2 such that

$$c_1 \cdot g(n) \leq f(n) \leq c_2 \cdot g(n)$$



Example :- $f(n) = 4 \cdot n^3 + 10 \cdot n^2 + 5n + 1$

$$g(n) = n^3, 4g(n) \leq f(n) \leq 5g(n)$$

Complexity of $f(n) = O(g(n))$
i.e $O(n^3)$.

Little Oh Notation (Θ)

Θ notation is used to denote an upper bound

$$\Theta(g(n)) = \left\{ \begin{array}{l} f(n) : \text{for any pos const. } C > 0 \\ \text{if a const } n_0 > 0 \end{array} \right.$$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0 \quad \text{such that } 0 \leq f(n) \leq Cg(n) \quad \forall n \geq n_0$$

Example :- $f(n) = n^4 + 35n^2 + 84$

$$g(n) = n^4$$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \lim_{n \rightarrow \infty} \frac{n^4 + 35n^2 + 84}{n^4}$$

$$\lim_{n \rightarrow \infty} \frac{n^4}{n^4} + \lim_{n \rightarrow \infty} \frac{35n^2}{n^4} + \lim_{n \rightarrow \infty} \frac{84}{n^4}$$

$$1 + 0 + 0 = 1$$

$$f(n) \neq \Theta(n^4)$$

Little Omega (ω)

If is denoted by ω . If the function $f(n) = \omega(g(n))$

$$\text{iff } \lim_{n \rightarrow \infty} \frac{g(n)}{f(n)} = 0$$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \infty$$

It defines the lower bound of Algo. which is not asymptotically tight

Divide & Conquer Approach

In divide and conquer approach, a problem is divided into smaller problems, then the smaller problems are solved independently, and finally the solutions of smaller problems are combined into a solution for the large problem.

Generally divide & conquer algorithms have 3 parts:-

- (*) Divide the problem :- Divide the problem into a number of sub problems that are smaller instances of the same problem.
- (*) Conquer the sub-problems :- Conquer the sub problems by solving them recursively. If they are small enough solve the sub-problems as base cases.
- (*) Combine the solutions :- Combine the sub-problems into the solution for the original problem.

Application of Divide and Conquer Approach

Following are some problems, which are solved using divide & conquer approach :-

- (*) Finding the Maximum & minimum of a sequence of numbers.
- (*) Strassen's Matrix Multiplication
- (*) Merge Sort.
- (*) Binary Search.

MERGE SORT

- Merge Means Combining
- Sorting means arranging Elements in increasing or decreasing order.
- Merge Sort has ~~the~~ operations :-
 - I) Divide
 - II) Conquer
 - III) Merge

Divide :- Divide the n-elements sequence to be sorted into two subsequences of $n/2$ Elements each.

Conquer :- Sort the subsequences recursively using Merge Sort.

Combine :- Merge the two sorted subsequences to produce the sorted answer.

Algorithm

A [0.....n-1], low, mid, high]

|| input :- Array A of unsorted Elements, low as beginning of pointer Array A and high as End of pointer Array A.

|| output :- Sorted Array A[0....n-1]

If (low < high) then

$$\text{mid} \leftarrow (\text{low} + \text{high})/2$$

Merge Sort (A, low, mid)

Merge Sort (A, mid+1, high)

Sorted Merge Sort (A, mid, low, high)

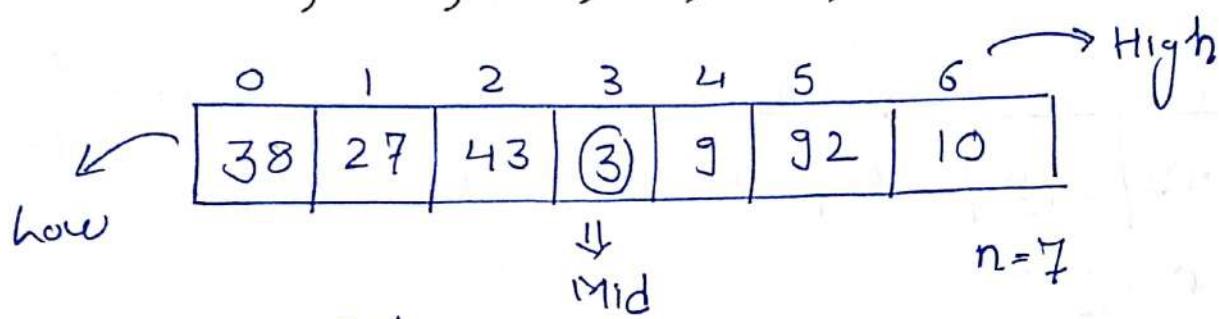
|| left sublist

|| right sublist

(A, mid, low, high)

Sort the following using Merge Sort :-

38, 27, 43, 3, 9, 92, 10



0 to n-1

0 to 7-1

0 to 6

Now $(0+6)/2 = 3$ Mid

Iteration 1

38	27	43	3	9	92	10
----	----	----	---	---	----	----

38	27	43	3
----	----	----	---

9	92	10
---	----	----

Iteration 2

38	27
----	----

43	3
----	---

9	92
---	----

10

38	27
----	----

43	3
----	---

9	92
---	----

10

27	38
----	----

3	43
---	----

9	92
---	----

10

Iteration 5

3	27	38	43
---	----	----	----

9	10	92
---	----	----

3	9	10	27	38	43	92
---	---	----	----	----	----	----

Sorted Elements

Sort the following using Merge Sort.

286, 45, 278, 368, 475, 389, 656, 788, 503, 126

0	1	2	3	④	5	6	7	8	9
286	45	278	368	475	389	656	788	503	126

$$mid = (0+9)/2 = 4.5$$

Iteration ①

0	1	2	3	4
286	45	278	368	475

Iteration ②

0	1	2
286	45	278

0	1
368	475

Iteration ③

286	45
278	

278

368	475

Iteration ④

286	45
45	286

278

368	475

Iteration ⑤

45	286
278	

278

368	475

Iteration ⑥

45	286
278	

278

368	475

Iteration ⑦

45	278	286
368		

368	475

389	656	788
389	656	788

0	1	2	3	4
389	656	788	503	126

0	1	2
389	656	788

0	1
503	126

389	656	788
389	656	788

503	126
503	126

389	656	788
389	656	788

126	503
126	503

126	389	503	656	788
126	389	503	656	788

126	389	503	656	788
126	389	503	656	788

Sorted

Array

45	126	278	286	368	389	475	503	656	788
----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Time Complexity

For Each algorithm we can calculate the time complexity by using three methods —

- (i) Master Method
- (ii) Tree Method
- (iii) Substitution Method.

Substitution Method

We make a guess for the soln & then we use Mathematical induction to prove the guess is correct or incorrect

Tree Method

In this Method we draw a recurrence tree & calculate the time taken by every level of tree. Finally we sum the work done at all levels. To draw the recurrence tree, we start from the given recurrence and keep drawing till we find a pattern among levels. The pattern is typically a arithmetic or geometric series.

Master Method

Master Method is a direct way to get the solution. The Master method works only for the following type of recurrences or for the recurrences that can be transformed to the following type.

$$T(n) = aT(n/b) + f(n) \quad \text{where } a > 1 \\ b > 1$$

Time Complexity For Merge Sort.

$$T(n) = T\left(\frac{n}{2}\right) + T\left(\frac{n}{2}\right) + Cn$$

occurrence
 Relation ↓ Time taken
 for Left sublist for Right sublist Time taken
 Merge left &
 Right sublist

$$T(n) = 2T\left(\frac{n}{2}\right) + Cn \quad \text{--- (1)}$$

$$T(n) = 2T\left(\frac{n}{2}\right) + Cn$$

$$\boxed{\text{Substitute } n = 2^k}$$

$$T(2^k) = 2T\left(\frac{2^k}{2}\right) + C \cdot 2^k \quad \text{--- (2)}$$

$$= 2T(2^{k-1}) + C \cdot 2^k$$

$$\boxed{\text{If we put } k = k-1}$$

$$2[2T(2^{k-2}) + C \cdot 2^{k-1}] + C \cdot 2^k$$

$$4T(2^{k-2}) + 2 \cdot C \cdot 2^{k-1} + C \cdot 2^k$$

$$4T(2^{k-2}) + 2 \cdot C \cdot 2^k / 2 + C \cdot 2^k$$

$$4T(2^{k-2}) + C \cdot 2^k + C \cdot 2^k$$

$$4T(2^{k-2}) + 2 \cdot C \cdot 2^k$$

$$2^2 T(2^{k-2}) + 2 \cdot C \cdot 2^k$$

$$2^2 T(2^{k-2}) + 2 \cdot C \cdot 2^k$$

$$2^3 T(2^{k-3}) + 3 \cdot C \cdot 2^k$$

$$2^4 T(2^{k-4}) + 4 \cdot C \cdot 2^k$$

$$2^k \cdot T(2^0) + k \cdot C \cdot 2^k$$

$$2^k T(2^0) + k \cdot C \cdot 2^k$$

$$2^k T(1) + k \cdot C \cdot 2^k$$

$$2^k (0 + k \cdot C \cdot 2^k)$$

$$T(2^k) = k \cdot C \cdot 2^k \rightarrow ③$$

$n = 2^k$ logarithm on both sides :-

$$\boxed{\log_2 n = k}$$

Substitute $k = \log_2 n$ in $③$

$$\log_2 n \cdot n$$

Time Complexity for Merge Sort
 $O(n \log_2 n)$

QUICK SORT

Quick Sort is also based on the concept of Divide and Conquer, just like Merge Sort. But in Quick Sort all the heavy lifting (heavy work) is done while dividing the array into subarrays, while in case of Merge Sort, all the real work happens during Merging the subarrays. In case of Quick sort, the Combine step does absolutely nothing.

Divide :- Split the array Elements into two sub-arrays based on pivot Element. Each Element in the left array should be less than or equal than the pivot Element. Each Element in the Right subarray should be greater than the pivot Element.

Conquer :- Sort the Sub arrays.

Combine :- Combine the left subarray and right subarray Elements into a single group.

We will use the following rules in Quick Sort

- ① if $A[\text{pivot}] > A[i]$ increment i -position
- ② if $A[\text{pivot}] < A[j]$ decrement j -position
- ③ if i and j cannot move then swap $A[i]$ & $A[j]$
- ④ Swap $A[\text{pivot}]$ and $A[j]$ if ' i ' crosses ' j '

Example :- 44 33 11 55 77 90 40 60 99 22 88
 let 44 be the pivot element & scanning done from right to left

Comparing 44 to the right side elements. If right side elements are smaller than 44, then swap it. As 22 is smaller than 44 so swap them.

22	33	11	55	77	90	40	60	99	44	88
----	----	----	----	----	----	----	----	----	----	----

Now Comparing 44 from left side Element & the Element must be greater than 44 then swap them. As 55 are greater than 44 so swap them

22	33	11	44	77	90	40	60	99	55	88
----	----	----	----	----	----	----	----	----	----	----

Recursively repeating step ① & ② until we get two lists one left from pivot element 44 & one right from pivot element.

22	33	11	40	77	90	(44)	60	99	55	88
----	----	----	----	----	----	------	----	----	----	----

Swap with 77

22	33	11	40	(44)	90	77	60	99	55	88
----	----	----	----	------	----	----	----	----	----	----

Now the elements on the right side and left side are greater & smaller than 44.

⇒ Now we get two sorted lists

22	33	11	40	44	90	77	60	99	55	88
----	----	----	----	----	----	----	----	----	----	----

Sublist ①

Sublist ②

And these are sorted under the same process as above done.

22 33 11 40
 Pivot
 11 33 22 40
 11 22 33 40

44 90 77 60 99 55 88
 Pivot
 88 77 60 99 55 90
 88 77 60 90 55 99
 88 77 60 55 90 99
 Sublist 3 Sublist 4

55 77 60 88 90 99
 55 77 60
 55 60 77
 Sorted.

Now Merging all sublists

11	22	33	40	44	55	60	77	88	90	99
----	----	----	----	----	----	----	----	----	----	----

Quick Sort Worst Case Complexity :- $T(n) = n^2$

Quick Sort Average Case Complexity $T(n) = n \log n$

SHELL SORT

Shell Sort is a highly efficient sorting algorithm and is based on Insertion Sort algorithm.

This algorithm avoids large shifts as in case of Insertion Sort, if the smaller value is too far right has to be moved to the far right & has to be moved to the far left.

This algorithm uses Insertion Sort on a widely spaced Elements, first to sort them and then sorts the less widely spaced Elements.

This spacing is termed as interval.

The Shell Sort sometimes called the "Diminishing Increment Sort," improves on the Insertion Sort by breaking the original list into a Number of smaller sublists. Each of which sorted using an insertion sort.

Algorithm for Shell Sort

Step① Initialize the value of Elements

Step② Divide the list into smaller sub-list of equal interval

Step③ Sort those sublists using Insertion Sort

Step④ Repeat Until Complete list is sorted

Sort the following using Shell Sort

54, 26, 93, 17, 77, 31, 44, 55, 20

(54)	26	93	(17)	77	31	(44)	55	20
------	----	----	------	----	----	------	----	----

Sublist 1

54	(26)	93	17	(77)	31	44	(55)	20
----	------	----	----	------	----	----	------	----

Sublist 2

54	26	(93)	17	77	(31)	44	55	(20)
----	----	------	----	----	------	----	----	------

Sublist 3

A shell sort with increment of three

(54)	26	93	(44)	77	31	(54)	55	20
------	----	----	------	----	----	------	----	----

Sublist 1 sorted

54	(26)	93	7	(55)	31	44	(77)	20
----	------	----	---	------	----	----	------	----

Sublist 2 sorted

54	26	(20)	7	77	(31)	44	55	(93)
----	----	------	---	----	------	----	----	------

Sublist 3 sorted

17	26	20	44	55	31	54	77	93
----	----	----	----	----	----	----	----	----

A shell sort after sorting each sublist.

Now Applying Insertion sort:-

17	26	20	44	55	31	54	77	93
----	----	----	----	----	----	----	----	----

17	20	26	44	56	31	54	77	93
----	----	----	----	----	----	----	----	----

17	20	26	31	44	55	54	77	93
----	----	----	----	----	----	----	----	----

17	20	26	31	44	54	55	77	93
----	----	----	----	----	----	----	----	----

Recurrence Equation

Recurrence Equation is an Equation that defines Sequence Recursively.

It is normally in the form of :-

$$T(n) = T(n-1) + n \quad \text{--- } ①$$

$$T(0) = 0 \quad \text{--- } ②$$

Equation ① is called Recurrence Relation

Equation ② is called initial condition

Solving Recurrence Equation

There are three Methods to solve Recurrence Eqⁿ :-

(i) Substitution Method

(ii) Trace Method

(iii) Master Method

Substitution Method

In this Method of substitution a guess for the solution is made :-

~~Forward~~ → Forward } Substitution method.
 → Backward }

Forward Substitution

This method make use of initial condition in the initial term & value for the next is generated.

This process continued until some formula is guessed.

Backward Substitution

In this Method backward values are substituted Recursively in order to derive a formula

Example of Forward Substitution

$$T(n) = T(n-1) + n$$

$$\boxed{n=1} \quad T(1) = T(1-1) + 1 \quad n > 0$$
$$= T(0) + 1$$

$$T(1) = 0 + 1 = 1$$

$$\boxed{n=2} \quad T(2) = T(2-1) + 2$$
$$= T(1) + 2$$
$$= 1 + 2$$
$$T(2) = 3$$

$$\boxed{n=3} \quad T(3) = T(3-1) + 3$$
$$= T(2) + 3$$
$$= 3 + 3$$

$$T(3) = 6$$

Formed formula

$$\Rightarrow n\left(\frac{n+1}{2}\right) + n^2/2 + n/2$$

Complexity :- $O(n^2)$

Example of Backward Substitution

$$T(n) = T(n-1) + n \quad \text{--- (1)}$$

↓

$$n = n-1$$

$$T(n-1) = T(n-1-1) + n-1$$

$$T(n-1) = T(n-2) + n-1 \quad \text{--- (2)}$$

$$T(n) = T(n-2) + n-1 + n \quad \text{--- (3)}$$

$$n = n-2 \text{ in } \mathcal{E}q^n \quad ①$$

$$= T(n-2-1) + n - 2$$

$$T(n-2) = T(n-3) + n - 2$$

$$T(n) = T(n-3) + n - 2 + n - 1 + n \quad \text{--- } ④$$

$$n = k$$

$$T(n) = T(n-k) + n - k + 1 + n - k + 2 + n$$

$$= T(k-k) + k - k + 1 + k - k + 2 + n$$

$$= T(0) + 1 + 2 + n$$

$$= 0 + \underbrace{1+2+n}$$

$$= n(n+1)/2$$

$$= n^2$$

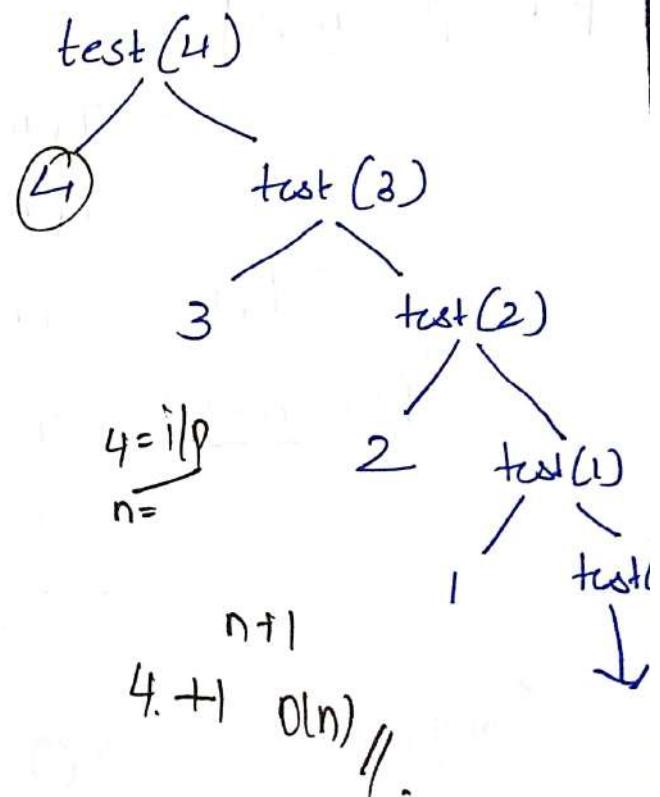
$$\text{Complexity} = O(n^2)$$

Tree Method

void test (int n)

```
{
    if (n > 0)
        {
            print ("d", n)
            test (n-1)
        }
}
```

```
}
```



Find Complexity for the following by Substitution Method.

$T(n) = \text{void test}(\text{int } n)$

$\left\{ \begin{array}{l} \text{if } (n > 0) \\ \quad \left\{ \begin{array}{l} \text{print} (" \%d ", n) \\ \quad \text{test}(n-1) \end{array} \right. \\ \quad \} \\ \quad \} \quad T(n) = T(n-2) + 2 \end{array} \right.$

Sol :- $T(0) = 1$

$$T(n) = T(\underbrace{n-1}) + 2 \rightarrow ①$$

put $n = n-1 \Rightarrow T(n-1-1) + 2$ $n = n-1$
 $T(n-1) = T(n-2) + 2$ $= \underbrace{T(n-2) + 2}_{= T(n-2) + 2}$

Substitute in ① $T(n) = T(n-2) + 2 + 2 = T(n-2) + 2 + 2$

$$T(n) = \underbrace{T(n-2) + 4}_{= T(n-2) + 4} \rightarrow ②$$

put $n = n-2$

$$\begin{aligned} &= T(n-2-1) + 2 \\ &= T(n-3) + 2 \end{aligned}$$

$$T(n) = T(n-3) + 6$$

put $n = k$ $= T(n-k) + k$

$$\begin{aligned} &\Rightarrow T(n-n) + n \\ &= T(0) + n \\ &= 1 + n \end{aligned}$$

Complexity $\Rightarrow O(n)$

Find the Complexity of below Recurrence.

$$T(n) = \begin{cases} 3T(n-1), & \text{if } n > 0 \\ 1 & \text{otherwise} \end{cases}$$

Let us solve using Recursion Substitution :-

$$\begin{aligned} T(n) &= 3T(n-1) \\ &= 3(3T(n-2)) \\ &= 3^2 T(n-2) \\ &= 3^3 T(n-3) \\ &= 3^4 T(n-4) \\ &\vdots \\ &= 3^n T(n-n) \\ &= 3^n T(0) \end{aligned}$$

This shows that the Complexity of the function
 $= 3^n$

Find recurrence eqn using substitution method :-

int recursive (int n)

{ if (n < 2)

return 1;

else

return recursive (n-1) + recursive (n-1)

$$T(n) = T(n-1) + T(n-1) + 1 + 1$$

$$T(n) = 2T(n-1) + C \quad \text{--- } ①$$

$$2 \{2T(n-2) + C\} + C \quad \text{--- } ②$$

$$T(n) = 2^2 T(n-2) + 3C$$

$$2^3 T(n-3) + 7C \quad \dots \quad 3^{nd}$$

$$T(n) = 2^3 T(n-3) + (2^3 - 1)C \quad \dots \quad 3^{nd}$$

$$T(n) = 2^K T(n-K) + (2^K - 1)C \quad \dots \quad K^{th}$$

$$\text{Hence } n-K = 1 \quad \{ (n-1) \leq 2 \}$$

$$K = n-1$$

Substitute K in (*)

$$T(n) = 2^{n-1} T(0) + (2^{n-1} - 1)C$$
$$2^{n-1} + (2^{n-1} - 1)C$$

$$O(2^{n-1})$$

$$O(2^n / 2)$$

$$\Rightarrow O(2^n)$$

HEAP SORT

J.N. J. Williams

This sorting Method discovered by
it consists of two stages :-

Heap → Heap Construction

→ Deletion of Minimum key

Heap Construction :- first we have to construct a heap for given Numbers.

Deletion of Minimum Key :- Delete Root key always ($n-1$)
to Remaining heap

for an array implementation of heap. Delete
the Element from the heap and push the
deleted Element in the last position of the array.
thus after deleting all the Elements from the heap
we will get the Elements in ascending or
descending order.

Heap

Heap is a Complete Binary tree or atmost
binary tree in which parent node be Either
greater or lesser than its Child Nodes.

Types of Heap

Max Heap

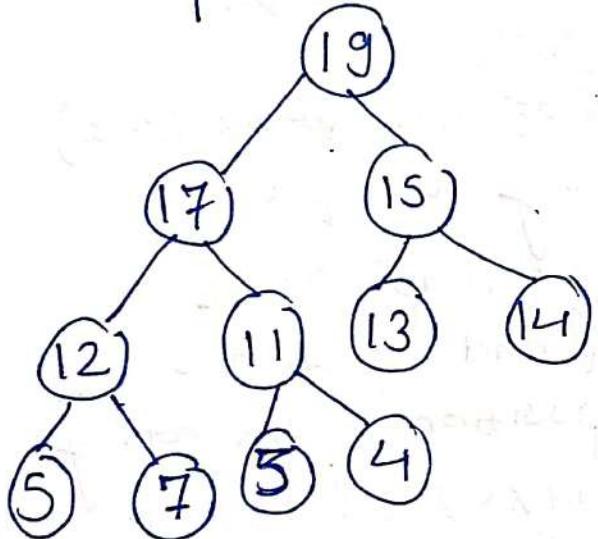
(Descending order)
Ascending

Min Heap

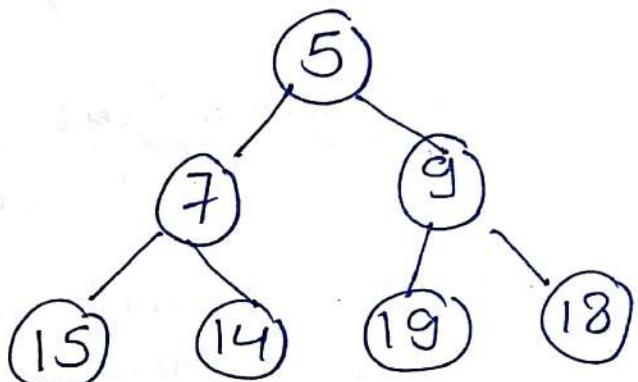
(Ascending order)
Descending

- * Max heap is a tree in which value of each node is smaller than the parent node.
 - * In heap the value of child node is greater than of parent node.
- Heap → [It should be complete binary tree
parented properly]

Maximum Heap



Minimum Heap



Algorithm for Max-heap

Maxheap : parent > child

for ($i = (n/2)$; $i > 0$; $i--$)

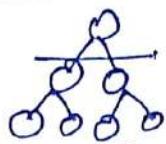
{

if takes $O(n/2)$ times

$\therefore O(n)$

Max Heap

Example:- 25, 57, 48, 37, 12, 92, 86, 33



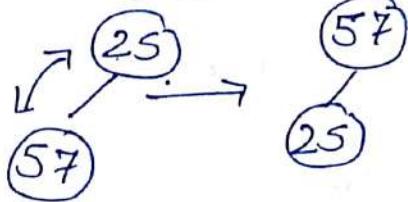
Binary tree



Step ①

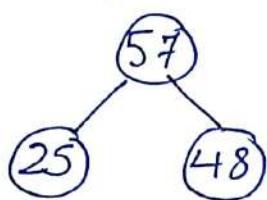


Step ②

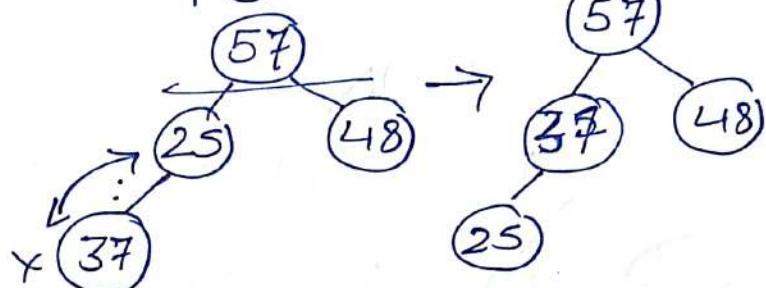


Max Heap
each value.
3 child nodes.

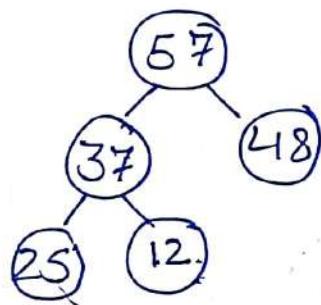
Step ③



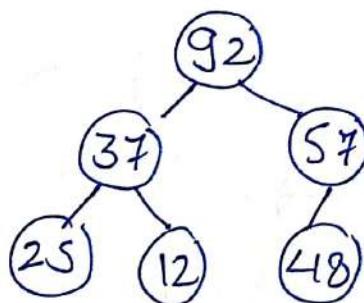
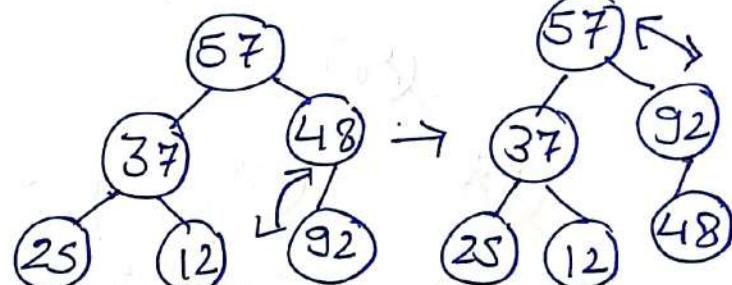
Step ④



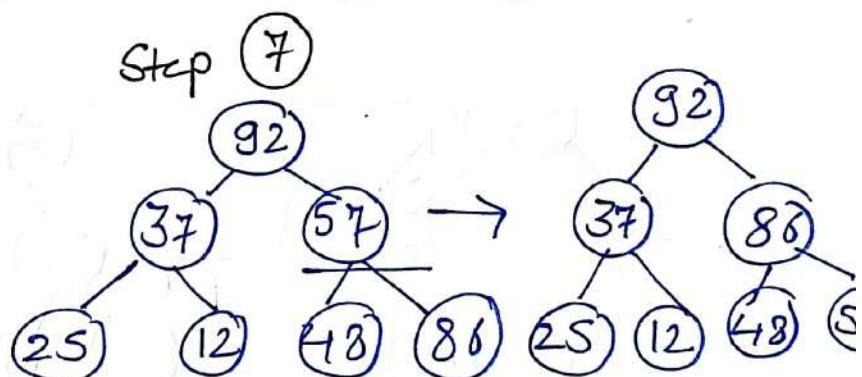
Step ⑤



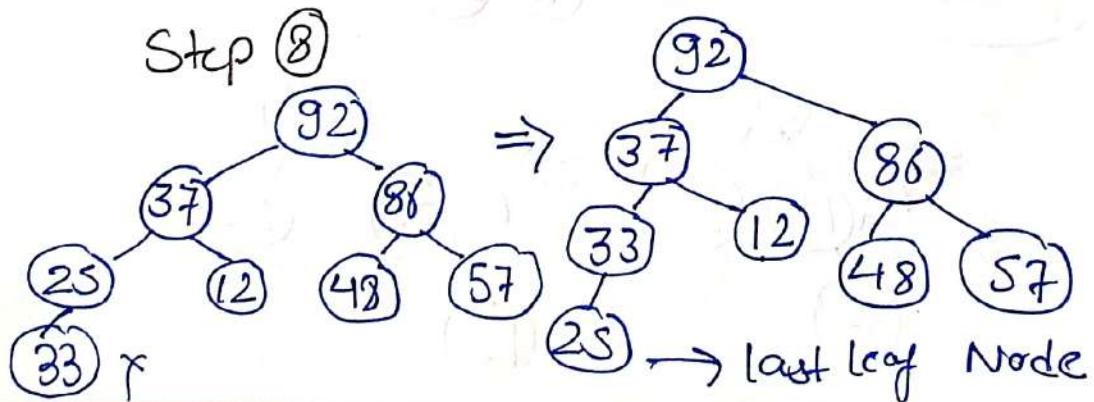
Step ⑥



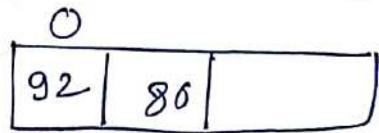
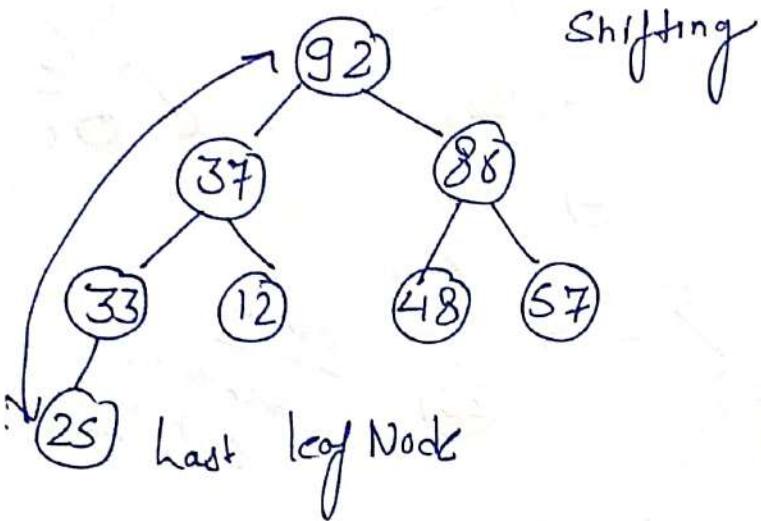
Step ⑦



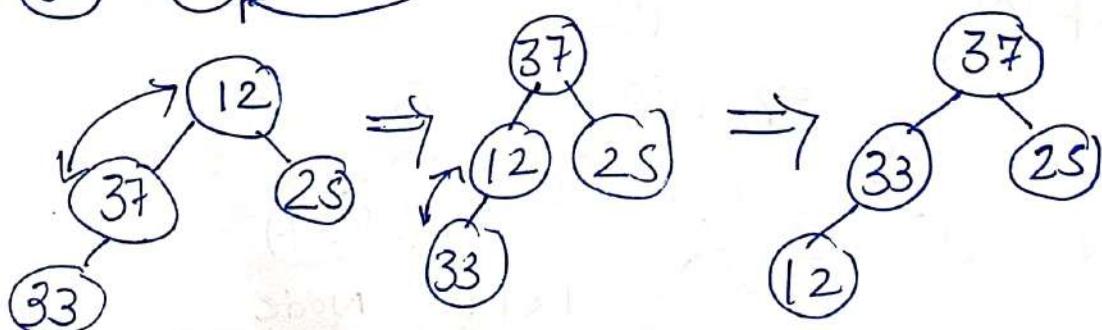
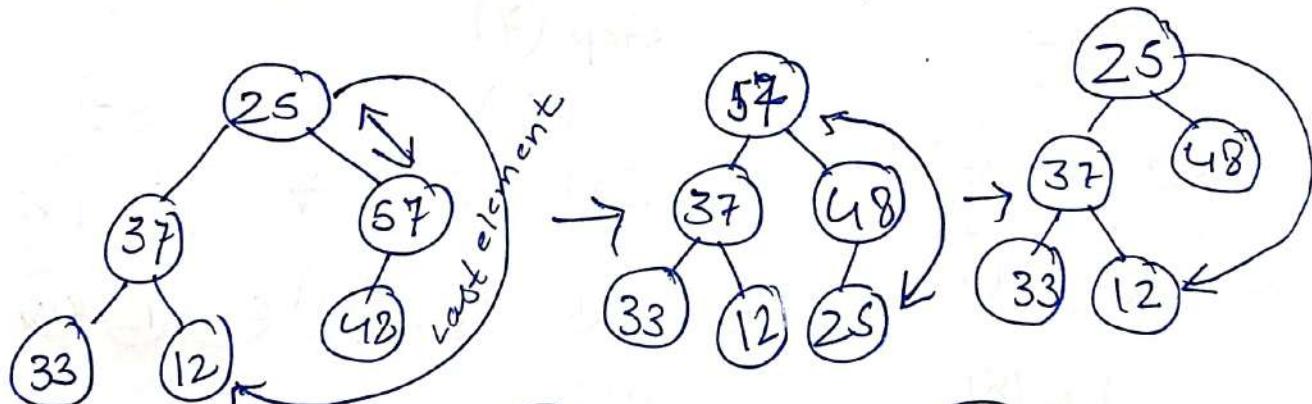
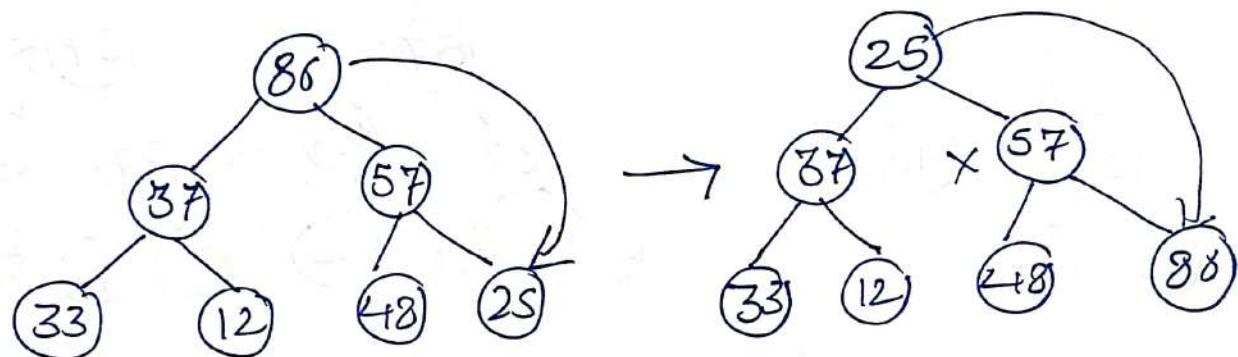
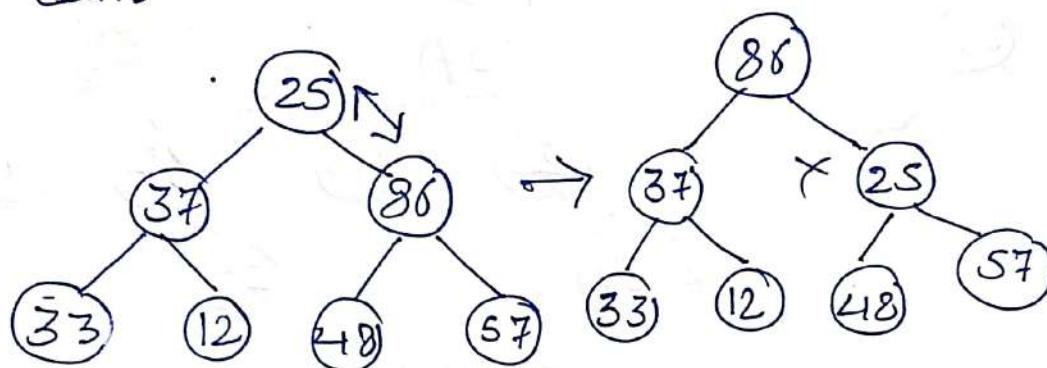
Step ⑧

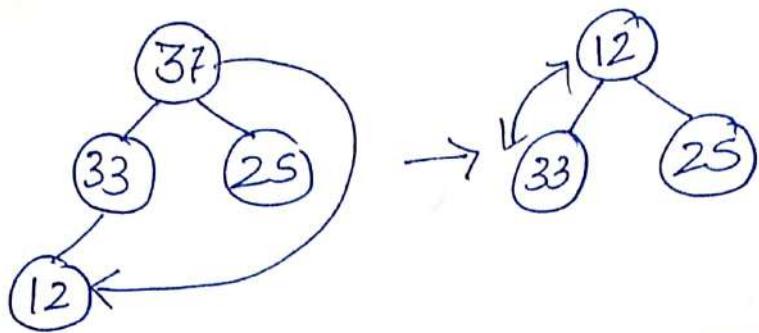


→ last leaf Node

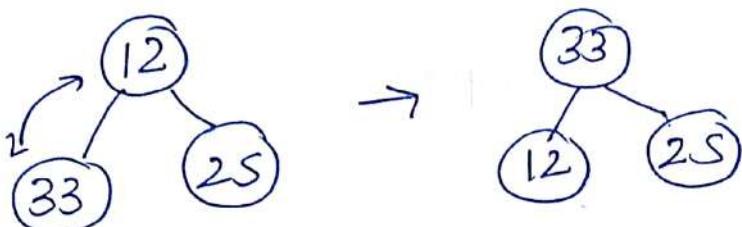


~~shifts~~

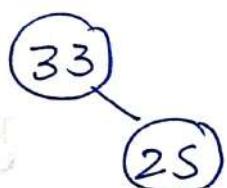




12	25	33	37	48	56	86	92
----	----	----	----	----	----	----	----



12	25	33	37	48	56	86	92
----	----	----	----	----	----	----	----



Sorted Array

12	25	33	37	48	56	86	92
----	----	----	----	----	----	----	----

Complexity of Heap Sort

Worst Case Time Complexity : $O(n * \log n)$

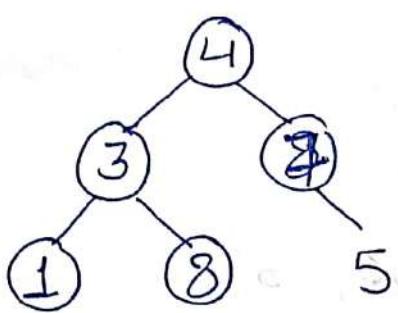
Best Case Time Complexity : $O(n * \log n)$

Average Time Complexity : $O(n * \log n)$

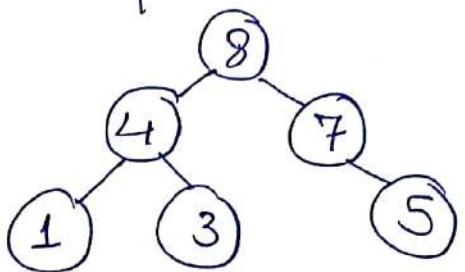
Space Complexity : $O(1)$

- * Heap Sort is not a stable sort and requires a constant space for sorting a list.

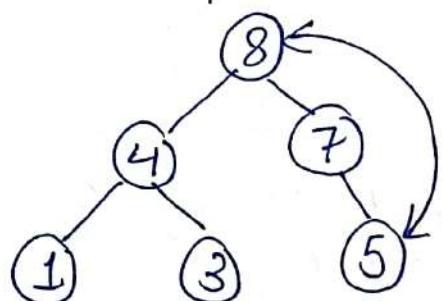
4, 3, 7, 1, 8, 5



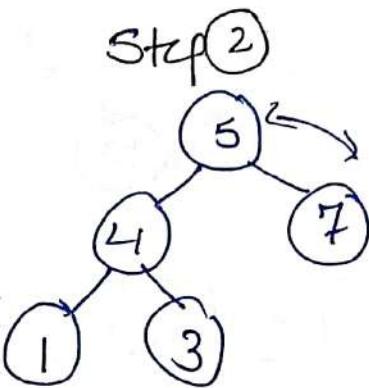
Max Heap



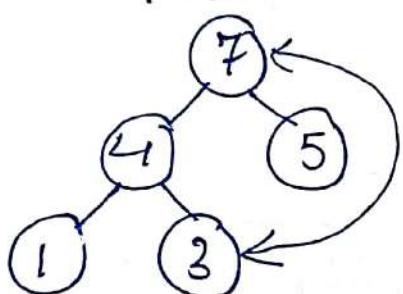
Step 1



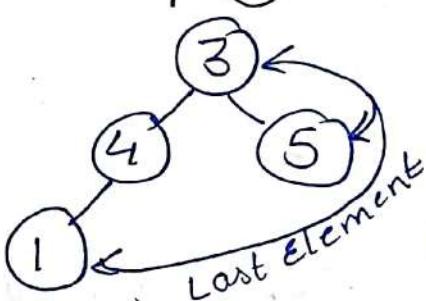
Step 2



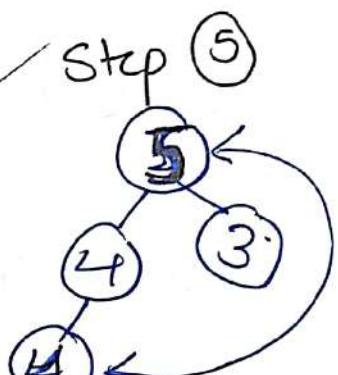
Step 3



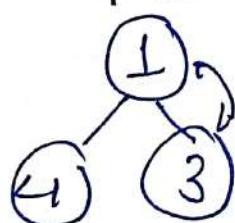
Step 4



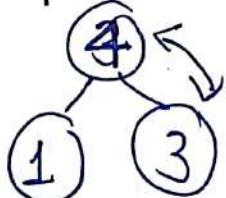
Step 5



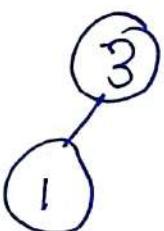
Step 6



Step 7



Step 8



Step 9



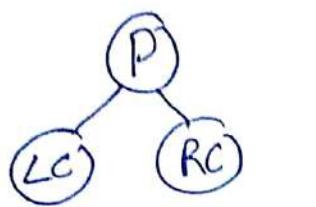
Step 10



1	3	4	5	7
---	---	---	---	---

Min-Heap Algorithm

Min-Heap Algorithm :- Decreasing order



$P \leq LC \& RC$

$P \rightarrow$ Parent Node

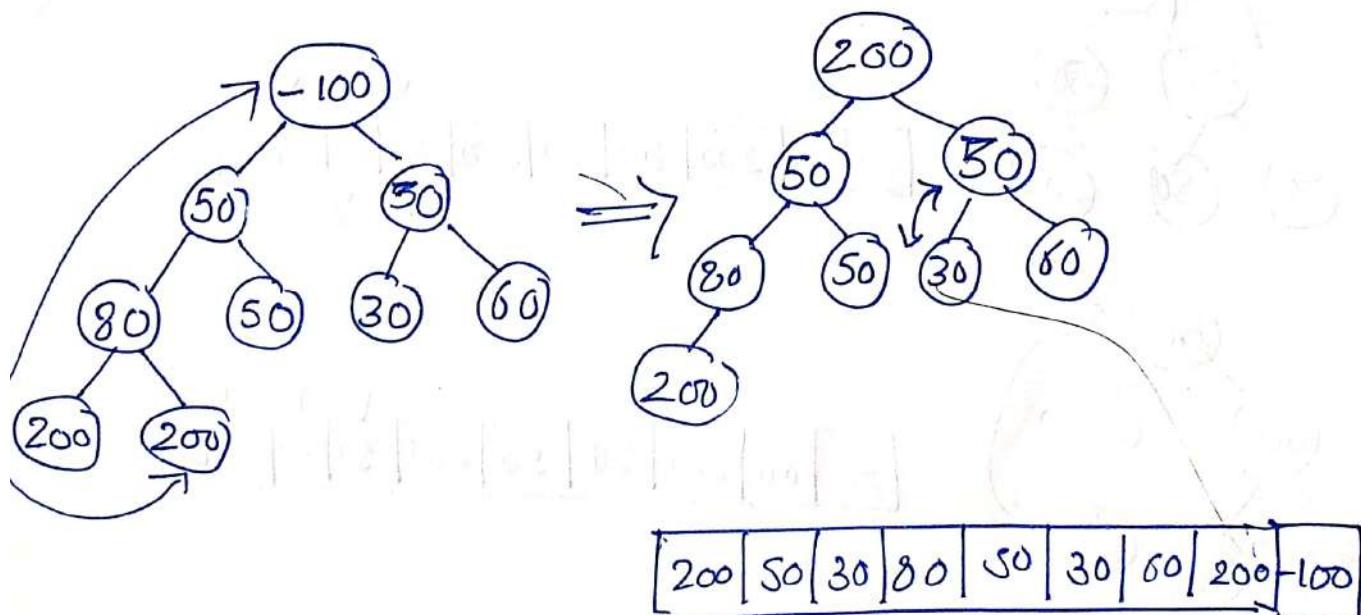
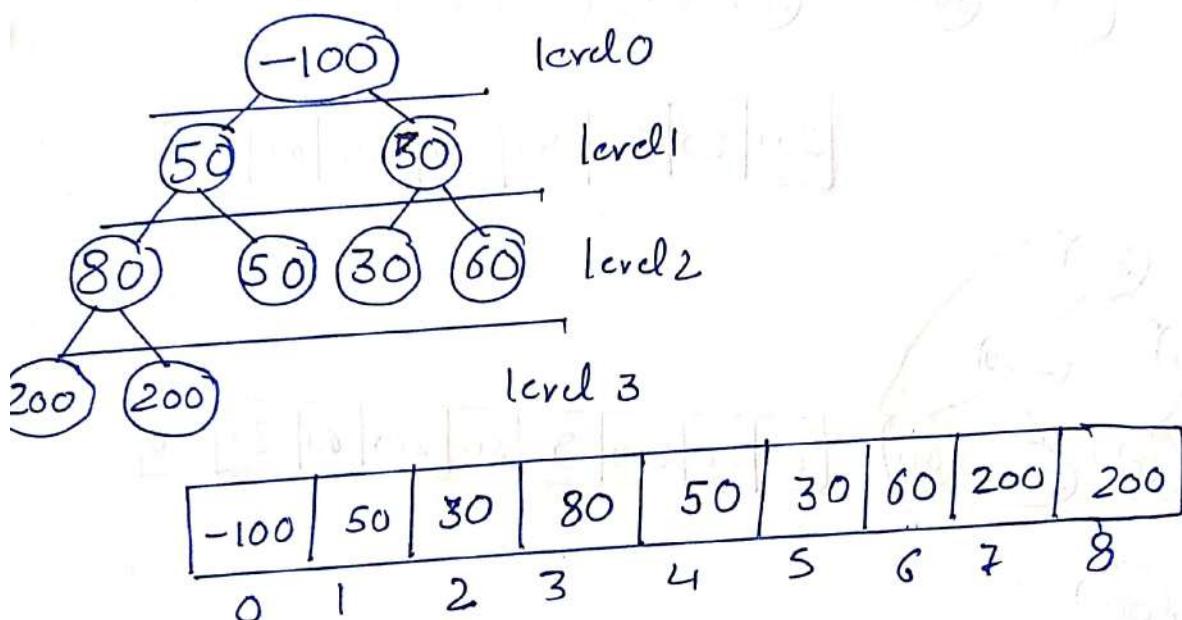
where LC - Left child

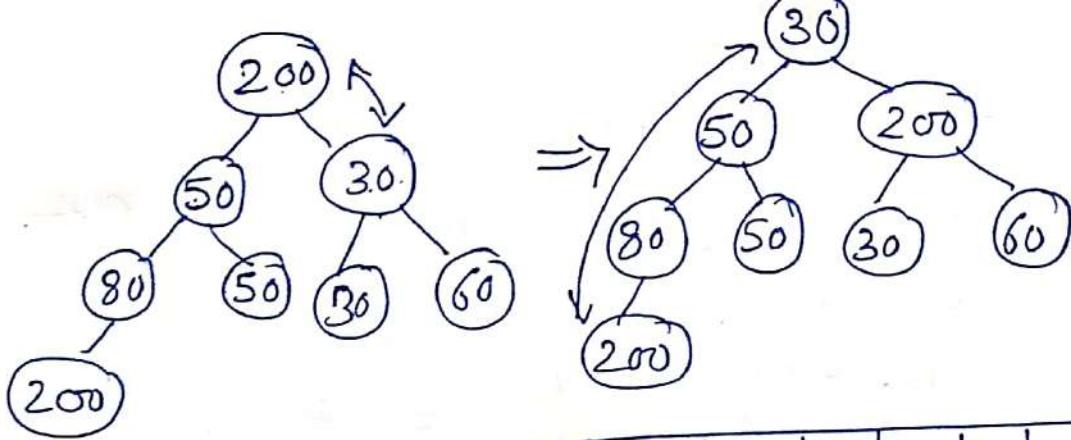
RC - Right child.

Example ①

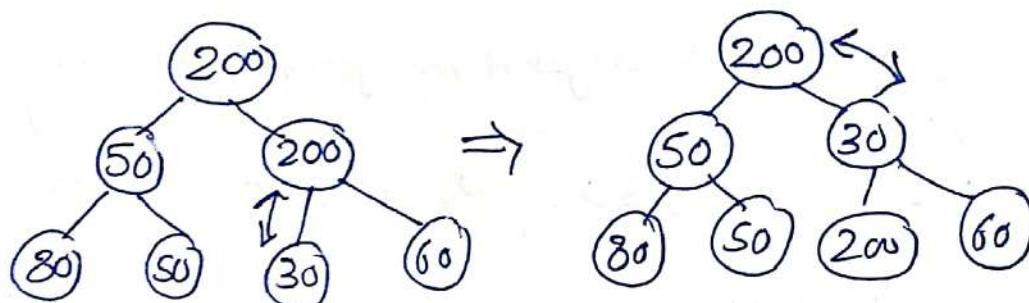
Show the trace of Heapsort algorithm for the following input data.

30, 50, -100, 200, 50, 30, 60, 80, 200 in order

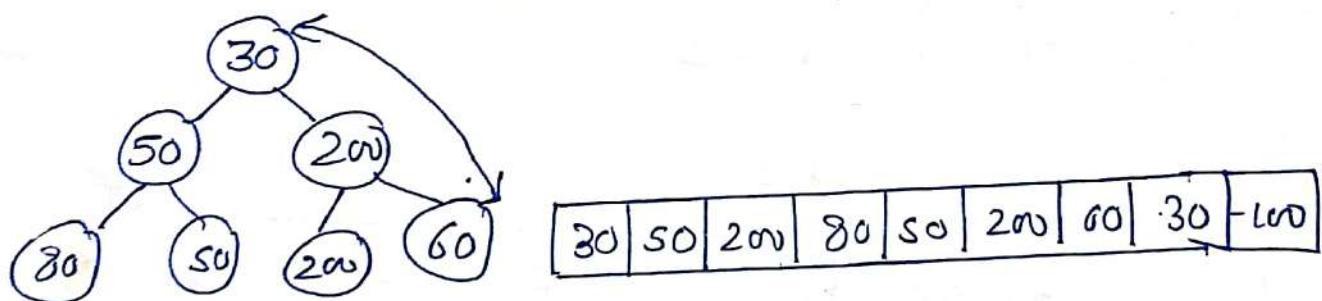




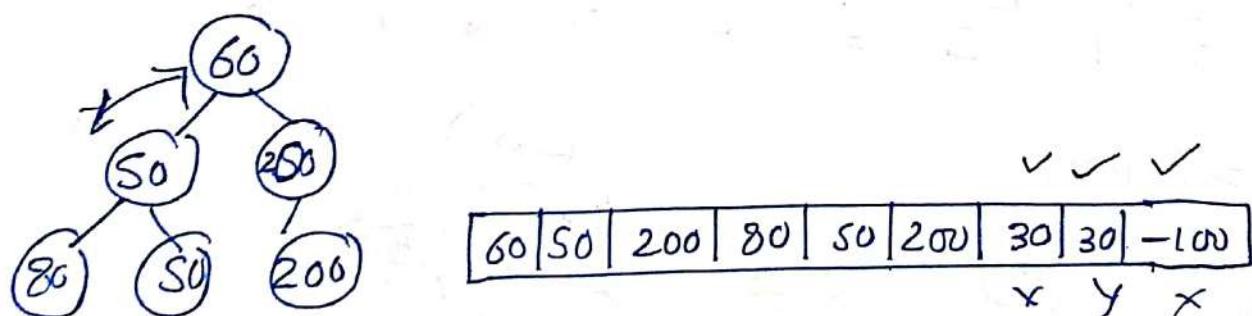
30	50	200	80	50	30	60	200	-100
----	----	-----	----	----	----	----	-----	------



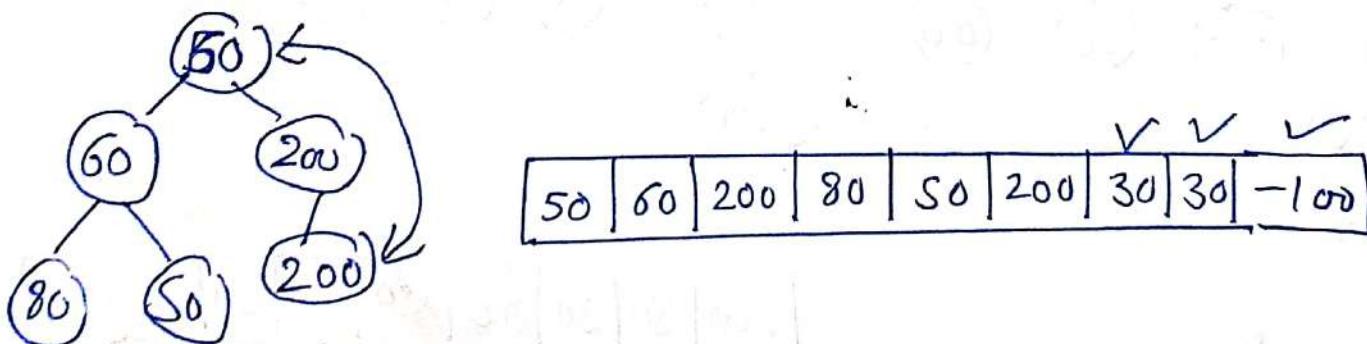
200	50	30	80	50	200	60	30	-100
-----	----	----	----	----	-----	----	----	------



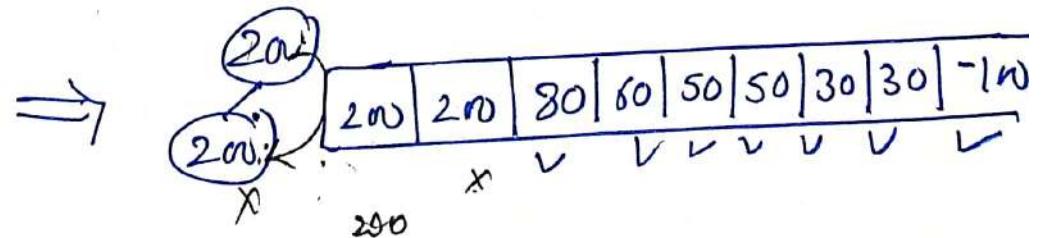
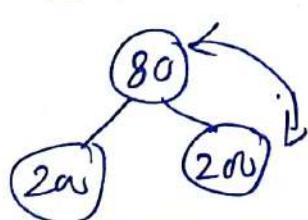
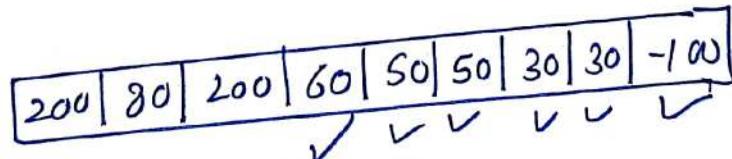
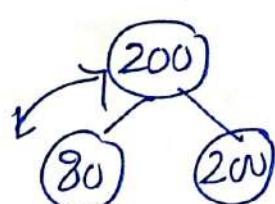
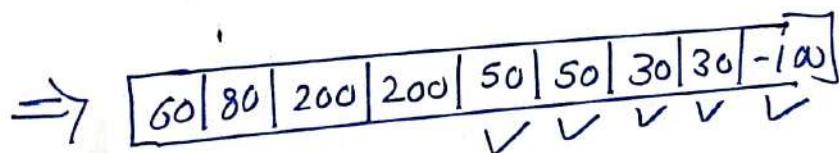
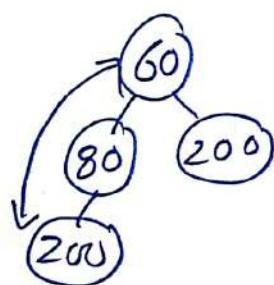
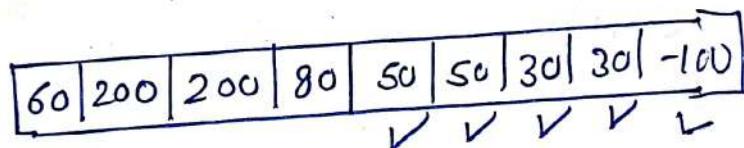
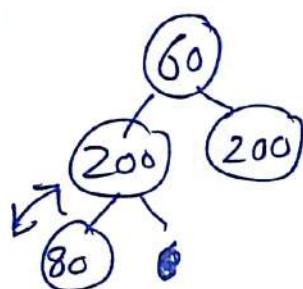
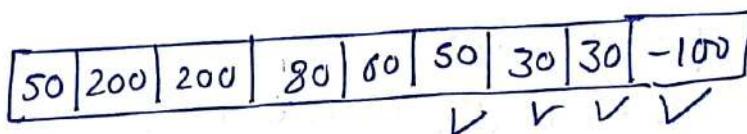
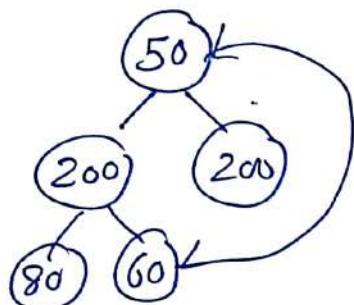
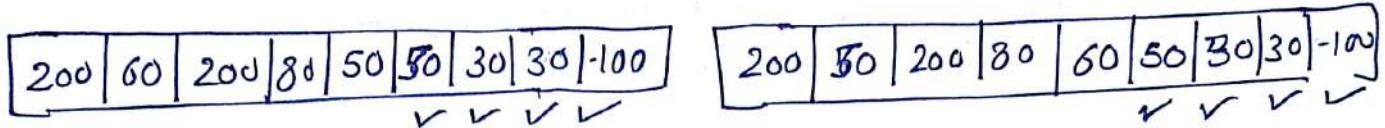
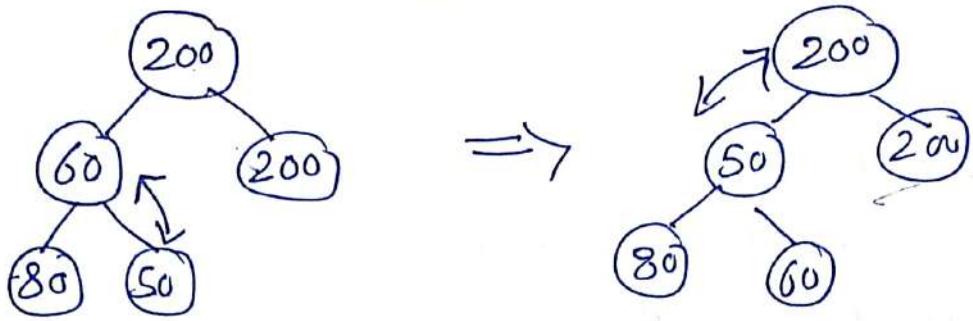
30	50	200	80	50	200	60	30	-100
----	----	-----	----	----	-----	----	----	------



60	50	200	80	50	200	30	30	-100
----	----	-----	----	----	-----	----	----	------



50	60	200	80	50	200	30	30	-100
----	----	-----	----	----	-----	----	----	------



Sorting

- * The Sorting problem is to rearrange the items of a given list in a specified order.
- * Sorting can be done on numbers, characters, strings or Records.
- * A sorting algorithm is called Stable if it preserves the relative order of any two equal elements in its input.
- * Sorting refers to the operation or technique of arranging and rearranging sets of data in some specific order. A collection of records, called a list where every record has one or more fields.

Categories of Sorting

Internal Sorting :- If all the data that is to be sorted can be adjusted at a time in the main Memory, the internal sorting Method is being performed.

External Sorting :- When the data is to be sorted cannot be accommodated in the Memory at the time & some has to be kept in Auxiliary Memory.

Serial Sorting :-

Bubble Sort :- A sorting algorithm which compares one Element to its next Element to its next Element & if requires it swaps like a bubble.

Selection Sort :- A sorting algorithm which selects a position in the Elements & compares it to the rest of the positions one by one.

Inscrition Sort :- It is the sorting mechanism where the sorted array is built having one item at a time.

Quick Sort :- A sorting algorithm which divides the elements into two subsets & again sorts recursively.

Heap Sort :- A sorting algorithm which is a comparison based sorting technique based on Binary Heap data structure.

Merge Sort :- A sorting algorithm which divides the elements to subgroups & then merges back to make sorted.

Radix Sort :- A sorting algorithm used for numbers. It sorted the elements by rank of the individual digits.

Comparison of Sorting techniques

Algorithm	Time Complexity			Extra space
	Best Case	Worst Case	Average Case	
Selection Sort	$O(n^2)$	$O(n^2)$	$O(n^2)$	1
Bubble Sort	$O(n)$	$O(n^2)$	$O(n^2)$	1
Inscrition Sort	$O(n)$	$O(n^2)$	$O(n^2)$	1
Merge Sort	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$	n (worst case)
Quick Sort	$O(n \log n)$	$O(n^2)$	$O(n \log n)$	$\log n$ (avg case), n (worst case)
Radix Sort	$O(s.d.n)$	$O(s.d.n)$	$O(s.d.n)$	n
Shell Sort	$O(n)$	$n \log^2 n$	$O(n \log^2 n)$ $O(n^{3/2})$	1

- (*) Different Sorting techniques have different uses according to their behaviour for different inputs. Every sorting technique has its own complexity & Quick sort is poorly sorted for humans.
- (*) For restricted data (data in a fixed interval) radix sort is useful.
- (*) Bubble sort is rarely used, but found in teaching & theoretical expressions.
- (*) Nowadays Shell sort is rarely used in serious applications. It does not use stack.
- (*) Some implementations of the Quick sort function in the C standard library targeted at Embedded Systems use it instead of Quick sort.
- (*) Selection Sort is useful where swapping is costly.
- (*) Insertion Sort is use for small data sets. Merge Sort & Quick sort are useful for large data sets.

Linear Time Sorting

- (*) There are sorting algorithms that run faster than $O(n \log n)$ time but they require special assumptions about the input sequence to be sorted.
- Example: of sorting algorithms that run in linear time are Counting sort, radix sort and bucket sort.
- (*) Despite of linear time usually these algorithms are not very desirable from practical point of view.

- * The Efficiency of Linear time algorithms depend on the keys randomly ordered. If this condition is not satisfied the result is the degrading in performance
- * These algorithms require Extra space proportional to the size of the array being sorted, so if we are dealing with large files.
- * The inner loop of these algorithms contain quite a few instructions so even though they are linear they would not be as faster than Quick sort