

Unit - IV

Searching & Sorting

Date : _____

Page : _____

Searching is of 2 types: (Based on type of data)

- + Internal
- + External. → (When data resides in sec memory & very vast.) & count of data

→ data is less, so main memory me

Types:
+ Linear search / sequential search.
+ Binary search.

Conditions:

- + Successful search
- + Unsuccessful search

Algo for linear search (P.T.O)

(1) if (`sea == a[i]`)

(2) + Set `POS = -1 & i = 0`
Repeat step (3) while `i <= N`,

if `A[i] == val`
set `POS = i`

Print `POS`.

Go to step (5).

Algo for Binary search:

Binary (A, LB, UB, Item, Loc, Mid, BEG, END)

Initialise
Set variable
`BEG = LB, END = UB.`

`MID = INT(BEG + END)/2.`

(1) → Repeat Step (3) & (4) while
`BEG <= END` & `A[Mid] != item`.

(2) if `item < A[Mid]`,
 then set `End = Mid - 1;`

(3) else Set `BEG = MID + 1.` (end of if)

(4) Set `Mid = int (BEG + END)/2.`
(end of step 2 & loop)

(5) if `A[Mid] == item` then set `Loc = Mid`,
 else set `Loc = Null.`
(end of if ~)

(6) exit.

Pass - 1.]

35	38	30	30	30	30
50	52	35	35	35	35
50	50	35	35	35	35
65	65	65	65	65	65
25	25	25	25	25	25
15	15	15	15	15	15
55	55	55	55	55	55
85	85	85	85	85	85

30 35 50 25 15 65 55 85.

30 35 50 25 15 55 65 85

Pass 3:-
30 35 50 25 15 55 65 85

30 35 50 25 15 55 65 85

30 35 25 80 15 55 65 85

30 35 25 15 50 55 65 85.

30 35 25 15 50 55 65 85

30 35 25 15 50 55 65 85

Pass 4:-
30 35 25 15 50 55 65 85
30 25 35 15 50 55 65 85.

Pass - 2

30 25 15 35 50 55 65 85.

30 25 15 35 50 55 65 85.

30 25 15 35 50 55 65 85.

30 25 15 35 50 55 65 85.

Pass 5:-

25 30 15 35 50 55 65 85

25 15 30 35 50 55 65 85

25 15 30 35 50 55 65 85.

25 15 30 35 50 55 65 85.

25 15 30 35 50 55 65 85.

25 15 30 35 50 55 65 85.

Pass 6:-

15 25 30 35 50 55 65 85.

15 25 30 35 50 55 65 85.

15 25 30 35 50 55 65 85.

15 25 30 35 50 55 65 85.

15 25 30 35 50 55 65 85.

15 25 30 35 50 55 65 85.

Pass 7 :-

15	25	30	35	50	55	65	85
15	25	30	35	50	55	65	85
15	25	30	35	50	55	65	85

15	25	30	35	50	55	65	85
15	25	30	35	50	55	65	85

15	25	30	35	50	55	65	85
15	25	30	35	50	55	65	85

32 51 27 85 66 23 13 57.

Pass 1 :-

32	51	27	85	66	23	13	57
32	27	51	85	66	23	13	57

32	27	51	85	66	23	13	57
----	----	----	----	----	----	----	----

32	27	51	66	85	23	13	57
----	----	----	----	----	----	----	----

32	27	51	66	23	85	13	57
----	----	----	----	----	----	----	----

32	27	51	66	23	13	85	57
----	----	----	----	----	----	----	----

32	27	51	66	23	13	57	85
----	----	----	----	----	----	----	----

Pass 2 :-

27	32	51	66	23	13	57	85
27	32	51	66	23	13	57	85
27	32	51	66	23	13	57	85

27	32	51	23	66	13	57	85
----	----	----	----	----	----	----	----

27 32 51 23 13 66 57 85

27 32 57 23 13 57 66 85

27 32 51 23 13 57 66 85

Pass 3 :-

27	32	51	23	13	57	66	85
27	32	51	23	13	57	66	85

27	32	23	51	13	57	66	85
----	----	----	----	----	----	----	----

27	32	23	13	51	57	66	85
----	----	----	----	----	----	----	----

27	32	23	13	51	57	66	85
----	----	----	----	----	----	----	----

Pass 4 :-

27	32	23	13	51	57	66	85
----	----	----	----	----	----	----	----

27	23	32	13	51	57	66	85
----	----	----	----	----	----	----	----

27	23	13	32	51	57	66	85
----	----	----	----	----	----	----	----

27	23	13	32	51	57	66	85
----	----	----	----	----	----	----	----

27	23	13	32	51	57	66	85
----	----	----	----	----	----	----	----

27	23	13	32	51	57	66	85
----	----	----	----	----	----	----	----

27	23	13	32	51	57	66	85
----	----	----	----	----	----	----	----

27	23	13	32	51	57	66	85
----	----	----	----	----	----	----	----

Pass 5:

27 23 13 32 57 57 66 85
23 27 13 32 57 57 66 85
23 13 27 32 57 57 66 85
23 13 27 32 57 57 66 85
23 13 27 32 57 57 66 85
23 13 27 32 57 57 66 85
23 13 27 32 57 57 66 85

Pass 6:

13 23 27 32 57 57 66 85
13 23 27 32 57 57 66 85
13 23 27 32 57 57 66 85

Algorithm:

Bubble Sort In, list, ptr, k

(1) Repeat step (2) & (3) for K=1 to N-1.

(2) Set PTR=1.

(3) Repeat while PTR ≤ N-K.

(a) if DATA[PTR] > DATA[PTR+1].

{ temp = DATA[PTR] = temp.
DATA[PTR] = DATA[PTR+1].
DATA[PTR+1] = temp.
} (end of if).

(b) Set PTR = PTR + 1.
} (end of inner loop).

~ end of step (1) loop.

(2) init.

Selection sort:

K=1 77 33 44 11 88 22 66 55
Loc 4 [1] [2] [3] [4] [5] [6] [7] [8]

K=2 11 33 44 77 88 22 66 55
Loc 6 +

K=3 11 22 44 77 88 33 55
Loc 6 +
K=4 11 22 33 77 88 44 66 55
Loc 6

K=5 11 22 33 44 88 77 66 55
Loc 8

$K=6$ 11 22 33 44 55 77 66 88
 $loc=7$

$K=7$ 11 22 33 44 55 66 77 88
 $loc=7$
→ 11 22 33 44 55 66 77 88

(1) $K=1$ 25 11 13 42 17 90 85 10 20
 $loc=8$

$K=2$ 7 11 13 42 17 90 85 25 10 20
 $loc=9$

$K=3$ 7 10 13 42 17 90 85 25 11 20
 $loc=9$

$K=4$ 7 10 11 42 17 90 85 25 13 20
 $loc=9$

$K=5$ 7 10 11 13 17 90 85 25 42 20
 $loc=5$

$K=6$ 7 10 11 13 17 90 85 25 42 20
 $loc=10$

7 10 11 13 17 20 85 25 42 90

$K=7$ 7 10 11 13 17 20 25 85 42 90
 $loc=8$

$K=8$ 7 10 11 13 17 20 25 85 42 90
 $loc=9$

$K=9$ 7 10 11 13 17 20 25 42 85 90
 $loc=9$

Algorithm for Selection Sort:

$Min(A, K, N, loc)$

(1) Set $min = A[k]$ & $loc = k$.

(2) Repeat for $j = k+1, k+2, k+3, k+4, \dots, N$.

if $min > A[j]$ & $loc = j$

$j = j + 1$
(end of loop).

(3) Return.

selection (A, N, \dots)

(1) Repeat step (2) & (3) for $k = 1, \dots, N-1$.

(2) Call $Min(A, K, N, loc)$.

(3) Interchange $A[k]$ & $A[loc]$.

$temp = A[k]$
 $A[k] = A[loc]$
 $A[loc] = temp$.

(end of step 1 to loop)

(4) Exit.

Insertion:

- (1) Set $A[0] = -1$.
- (2) Repeat step (3) to (5) for $k=1, 2, 3, \dots$
- (3) Set $\text{Temp} = A[k]$ & $\text{PTR} = k-1$.
- (4) Repeat while $\text{Temp} < A[\text{PTR}]$.
 - (a) Set $A[\text{PTR}+1] = A[\text{PTR}]$.
 - (b) Set $\text{PTR} = \text{PTR}-1$.

(end of loop).
- (5) Set $A[\text{PTR}+1] = \text{Temp}$.
(insert element at proper place)

[end of step 1 to loop]

- (6) Return.

Dry-run:

Pass $A[0]$ $A[1]$ $A[2]$ $A[3]$ $A[4]$ $A[5]$ $A[6]$ $A[7]$ $A[8]$.

1 -1 77 33 44 11 88 22 66 55

	$A[0]$	$A[1]$	$A[2]$	$A[3]$	$A[4]$	$A[5]$	$A[6]$	$A[7]$	$A[8]$
1	-1	77	33	44	11	88	22	66	55
2	-1	33	77	44	11	88	22	66	55
3	-1	33	44	77	11	88	22	66	55
4	-1	11	33	44	77	88	22	66	55
5	-1	11	33	44	77	88	22	66	55
6	-1	11	22	33	44	77	88	66	55
7	-1	11	22	33	44	88	77	88	55
8	-1	11	22	33	44	55	88	77	88
9	-1	7	3	2	10	5	08		
Pass	$A[0]$	$A[1]$	$A[2]$	$A[3]$	$A[4]$	$A[5]$	$A[6]$	$A[7]$	$A[8]$
1	-1	1	7	3	2	10	5	08	
2		1	7	3	2	10	5	08	
3		1	3	7	2	10	5	08	
4		1	2	3	7	10	5	08	
5		1	2	3	7	10	55	08	
6		1	2	3	5	7	10	08	
7		0	1	2	3	5	7	10	
8		0	1	2	3	5	7	8	
9		0	1	2	3	5	7	8	

Divide & compare method

Quick Sort :- (Appn of stack)

Quick(A, N, BEG, END, LEFT, RIGHT, LOC).

→ BEG & END contain boundary value of sublist of A to which this algorithm will be applied.

→ LOC keeps track of position of 1st element $A[BEG]$ of the sub-list during iteration.

→ Left & right contains the boundary value of the list of the element that has not been scanned.

(1) Set Left = BEG, Right = END & LOC = BEG.

(2) Scan from Right to left.

(a) Repeat while $A[LOC] < A[RIGHT]$ and $LOC \neq Right$.

Set Right = Right - 1.

End of loop.

(b) If $LOC = Right$ then Return.

(c) if $A[LOC] > A[RIGHT]$ then

(i) Interchange $A[LOC]$ & $A[RIGHT]$.

Temp = $A[LOC]$.

$A[LOC] = A[RIGHT]$.

$A[RIGHT] = Temp$.

(ii) Set $LOC = Right$.

(iii) Go to step (3).

(end of if)

(3) Scan from Left to Right.

(a) Repeat while $A[Left] \leq A[LOC]$ and $Left \neq Right$.

Set Left = Left + 1.

(end of loop).

(b) If $LOC = Left$ then return.

(c) if

c) if $A[\text{left}] > A[\text{loc}]$ then.

(i) Interchange $A[\text{left}]$ to $A[\text{loc}]$.

$$\begin{aligned} \text{temp} &= A[\text{loc}] \\ A[\text{loc}] &= A[\text{left}] \end{aligned}$$

$$A[\text{left}] = \text{temp}.$$

(ii) set $\text{LOC} = \text{left}$.

(iii) Go to step (2)
(end of if)
structure.

Algo - Quicksort (This algo sort array
with N element).

1) (Initialize) $\text{TOP} = \text{NULL}$

2) [Push boundary value A onto
the stack when A has
2 or more element].

if $N > 1$ then, $\text{TOP} = \text{TOP} + 1$.

$$\text{lower}[1] = 1$$

$$\text{Upper}[1] = N$$

(3) Repeat step 4 to 7 while $\text{TOP} \neq \text{NULL}$
(backside
now)

RADIX SORT. (more no. of data)

Eg. Roll No. list.

Google

Check from digit place

0
0 0
one's ten's
place place

Eg. 10, 15, 1, 60, 5, 100, 25, 50.

Step I :- find the largest no. i.e., to count
no. of digit in that no. if no.
is of n digit then no. of
pass = n .

II :- Convert all no.'s / input into n digit
no.

In this case $n = 3$ so,

No's. $\begin{array}{c} 010, 015, 001, 060, 005, 100, 025, 050 \\ \downarrow \text{unit} \\ \text{tens} \end{array}$

III :- Classify no. according to unit digit
place.

digit contain no's form.

0 - 9

assign label 0
1
2
3
4. ↓ smallest

These nos. contains 0 in unit place
assign them into table.

Queue for digit.

Pass I:
0 - 010, 060, 100, 050.
1 - 001.
2 - >
3 - >
4 - >
5 - 015, 005, 025.
6 - >
7 - >
8 - >
9 - >

Output of Pass-I

Starting with 1st queue or table:

010, 060, 100, 050, 001, 015, 005, 025

For pass II we've input, ie, output of Pass I.

Pass II: 010, 060, 100, 050, 001, 015, 005, 025
Check ten's place.

0 - 100, 001, 005
1 - 010, 015
2 - 025
3 - >
4 - >
5 - 050
6 - 060
7 - >
8 - >
9 - >

O/P of Pass II:

100, 001, 005, 010, 015, 025, 050, 060

This will be I/P for Pass III.

Pass III: 100, 001, 005, 010, 015, 025, 050, 060

0 - 001,005, 010, 015, 025, 050, 060
 1 - 100
 2 - ~
 3 - ~
 4 - x
 5 - y
 6 - >
 7 - >
 8 - x
 9 - >

Op - 001, 005, 010, 015, 025, 050, 060, 100.

Ques:- 348, 143, 361, 423, 538, 128, 321, 543, 366, 4281,
2163.

Ans:- largest no. = 4 digit (4281)

0348, 0143, 0361, 0423, 0538, 0128, 0321, 0543,
0366, 044281, 2163.

Pass I:- Check unit's digit!

0 - y
 1 - 0361, 0321, 4281
 2 - ~
 3 - 0143, 0423, 0543, 2163
 4 - p
 5 - ~

6 - 0366
 7 - ~
 8 - 0348, 0538, 0128
 9 - ~

O/p of Pass I:- 0361, 0321, 4281, 0143, 0423, 0543,
2163, 0366, 0348, 0538, 0128

Input for Pass II:-

0 - y
 1 - ~
 2 - 0321, 0423, 0128
 3 - 0538
 4 - 0143, 0543, 0348.
 5 - ~
 6 - 0361, 2163, 0366.
 7 - >
 8 - 4281
 9 - ~

Op:- 0321, 0423, 0128, 0538, 0143, 0543, 0348,
0361, 2163, 0366, 4281,

SIP for Pass II:- 0321, 0423, 0128, 0538, 0143, 0543,
0361, 2163, 0366, 4281,

Pass III:

0 - >
1 - 0128, 0143, 2163,
2 - 4281.
3 - 0321, 0348, 0361, 0366,
4 - 0423.
5 - 0538, 0543
6 - >
7 - >
8 - >
9 - >

o/p:-

0128, 0143, 2163, 4281, 0321, 0348, 0361,
0366, 0423, 0538, 0543

Pass IV:

0 - 0128, 0143, 0321, 0348, 0361, 0366, 0423,
1 - >
2 - 2163
3 - >
4 - 4281
5 - >
6 - >
7 - >
8 - >
9 - >

broke

o/p:- 0128, 0143, 0321, 0348, 0361, 0366, 0423,
0538, 0543, 2163, 4281.

ff Merge Sort:

1st make pair of 2 then apply sorting after that make group of 4 then apply sorting then group of 8 so on with power of 2.

Eg: 5, 3, 4, 6, 3, 1, 6, 2.

2, 5, 4, 6, 1, 3, 3, 6

2, 4, 5, 6, 1, 3, 3, 6

1, 2, 2, 3, 4, 5, 5, 6 - sorted list

Pairing

Eg: 20, 66, 33, 40, 20, 35, 88, 60, 11, 80, 20,

50, 44, 70

Pairing of 4:

33, 66, 20, 40, 35, 88, 11, 60, 20, 80,

44, 50, 70

Pairing of 8:

20, 33, 40, 66, 11, 35, 60, 88, 20, 44, 20, 80, 70

Pairing of 16:

11, 20, 33, 40, 55, 60, 66, 88, 20, 44, 50, 70, 77

11, 20, 20, 33, 40, 44, 50, 55, 60, 66, 70, 70, 77

Explanation of Quick Sort:

44 33 11 55 77 90 40 60 99 22 88 66
LB BEG Left Loc(fin)

44 ≤ 66 ✓

44 ≤ 88 ✓

44 ≤ 22 ✗ (swap).

22 33 11 55 77 90 40 60 99 44 88 66
↑ L.
22 ≤ 44 ✓ R (loc)
33 ≤ 44 ✓ fin
11 ≤ 44 ✓
55 ≤ 44 ✗ (swap).

22 33 11 44 77 90 40 60 99 55 88 66
↑ L
(fin)
(Loc).

44 ≤ 55 ✓

44 ≤ 99 ✓

44 ≤ 60 ✓

44 ≤ 40 ✗ (swap).

22 33 11 40 77 90 44 60 99 55 88 66
↑ L
R (fin)
(Loc)

44 44 ≤ 40.

40 ≤ 44 ✓

77 ≤ 44 ✗ (swap).

22 33 11 40 44 90 77 60 99 55 88 66
↑ L
R
(fin)
(Loc)

44 ≤ 77. ✓

44 ≤ 90 90 < 77. ✗ (swap)

22 33 11 40 44 77 90 60 99 55 88 66
↑ R
22 33 11 40 [44] 90 77 60 99 55 88 66
↑ L P R
ie, loc=right so, loop terminates

90 77 60 99 55 88 66.
↑
BEG
Left
LOC (fin)

$90 \leq 66$ ✓ (swap).

66 77 60 99 55 88 90.
↑
Left
 $66 \leq 90$ ✓ (LOC).

$77 \leq 90$ ✓

$60 \leq 90$ ✓

$99 \leq 90$ ✗ (swap).

66 77 60 90 55 88 99
↑
L
(fin)
LOC

$90 \leq 99$ ✓

~~90~~
 $90 \leq 88$ ✗ (swap)

66 77 60 88 55 90 99.
↑
L
LOC
 $88 \leq 90$. ✓

55 77 60 88 55 90 99.
↑
Right
LOC
(fin)

$66 \leq 55$ ✗ (swap).

55 77 60 88 66.
↑
L
(fin).

$55 \leq 66$ ✓
 $77 \leq 66$ ✗ (swap).

55 66 60 88 77.
↑
L
LOC (fin)

$66 \leq 77$ ✓.

$66 \leq 88$. ✓
 $66 \leq 60$ ✗ (swap).

55 60 ~~66~~ 88 77.
↑
L
LOC
(fin).

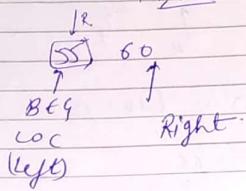
$60 \leq 66$.

77 88 77. 88 77. 88 77.
↑ ↑ ↑
L R R
LOC (fin) Swap
 $77 \leq 88$.

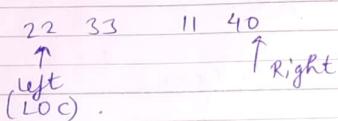
88 77. 88 77.
↑ ↑
L R
LOC (fin) Swap
 $77 \leq 88$.

89-fin.

77 88



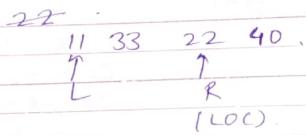
55 ≤ 60 ✓



11 ≤ 22 ✓

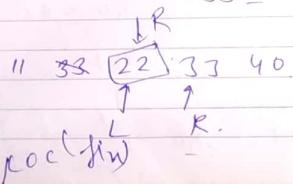
22 ≤ 40 ✓

22 22 ≤ 11 ✗ (swap)

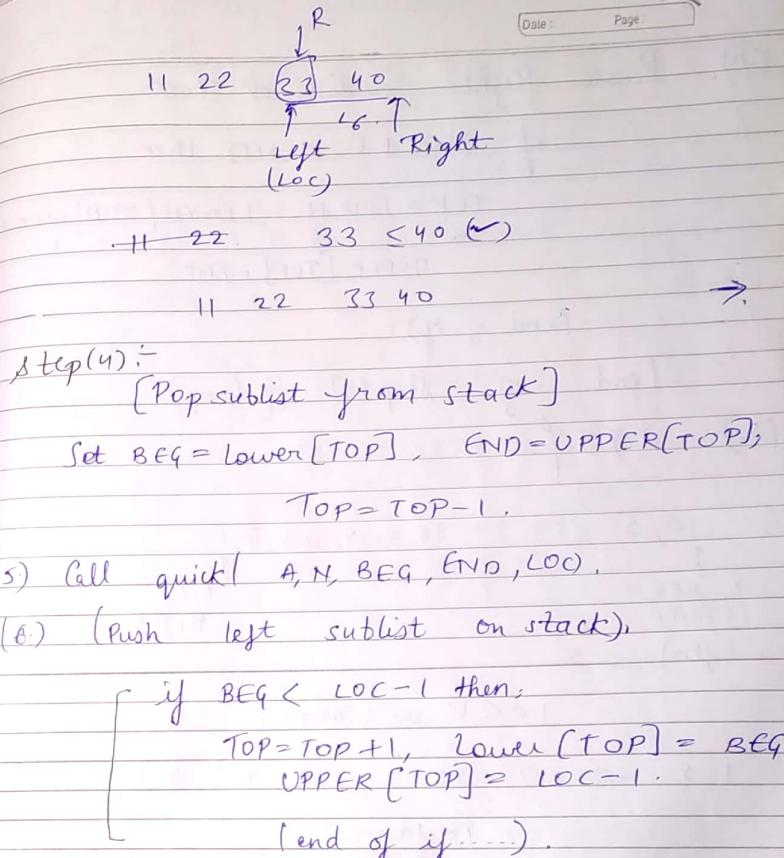


11 ≤ 22 ✓

33 ≤ 22 ✗ (swap)



22 ≤ 33 ✓



(A) (Push left sublist on stack),

{ if $BEG < LOC - 1$ then:

$TOP = TOP + 1$, $Lower[TOP] = BEG$,
 $UPPER[TOP] = LOC - 1$.

{end of if---).

to develop efficiency - sorted board.

) Push Right sublist on stack.

if $LOC + 1 < END$ then

$TOP = TOP + 1$, $LOWER[TOP] = LOC + 1$

$UPPER[TOP] = END$.

(end of if).

(end of step (3) & loop).

Ex If.

10, 5, 63, 27, 88, 3, 15, 43, 4, 1.
↑
BEG
(Left)
(fin)

END
Right.

$10 < 1$. \cancel{x} (swap?)

1 5 63 27 88 3 15 4 3 4 10.
 ~~$10 < 4$~~
 ~~$1 < 4$~~
 ~~$1 < 3$~~
 ~~$1 < 1$~~ .

$10 < 3$

$1 < 10 \checkmark$
 $5 < 10 \checkmark$
 $63 < 10 \cancel{x}$ (swap)

1 5 10 27 88 3 15 43 4 63
↑
LOC
Right.

$10 < 5$

$10 < 63 \checkmark$
 ~~$10 < 4$~~ \cancel{x} (swap).

1 5 4 27 88 3 15 43 10 63
↑
left
LOC
fin

$4 < 10 \checkmark$
 $27 < 10 \cancel{x}$ (swap).

1 5 4 10 88 3 15 43 27 63
↑
LOC
Right

$10 < 27 \checkmark$
 $10 < 43 \checkmark$
 $10 < 15 \checkmark$
 $10 < 3 \cancel{x}$ (swap).

1 5 4 3 88 10 15 43 27 63
↑
left
LOC
Right

$3 < 10 \checkmark$
 $88 < 10 \cancel{x}$ (swap).

15 43 10 88 15 43 27 63.
 Loc R 10 < 88 ✓.

88 15 43 27 63.
 Loc END.
 88 < 63 ✓ (swap).

63 15 43 27 (88).
 Left Loc.

63 < 88 ✓
 15 < 88 ✓
 63 < 27 ✓ (swap). 15 < 88 ✓
 27 < 88 ✓.

63 15 27 43 88.
 Loc Right 27 < 43 ✓.

63 15 43 27.
 Loc Right
 63 < 27 ✓ (swap).

27 15 43 (63).
 Left Loc.
 27 < 63 ✓
 15 < 63 ✓
 43 < 63 ✓.

27 15 43
 ↑ ↑
 Left Loc Right
 27 < 43 ✓
 27 < 15 ✓ (swap).

15 (27) 42.
 ↑ ↑
 Left Loc.
 15 < 27 ✓.

15 43.
 Loc Right

1 < 3. ✓
 1 < 4 ✓
 1 < 5 ✓

5 4 3.
 Loc Right
 5 < 3 ✓ (swap).

3 4 (5).
 Left Loc.
 3 < 5 ✓.
 4 < 5 ✓.

(3) 4.
 Loc Right
 3 < 4 ✓.

Gap is used to create sublist.

Shell Sort (Extension of Insertion Sort).

Compare elements by considering gap where $\text{gap} < \text{END}$.

Reduce value of gap with each iteration till $\text{gap} = 1$.

$$\text{for 1st iteration, } \frac{\text{gap}}{2} = \text{Floor} \left(\frac{n}{2} \right) \\ = \frac{9}{2} = 4.5$$

Pass 1: $\text{floor gap} = 4$.

Sub1 - $a[0], a[4], a[8]$

Sub2 - $a[1], a[5]$

Sub3 - $a[2], a[6]$

Sub4 - $a[3], a[7]$.

0	1	2	3	4	5	6	7	8
14	18	19	37	23	40	29	30	11
14				23		11		
18					40			
19					29			
						30		
							37	

-1 | 14, 23 | 11.

-1 | 14 23 | 11

-1 | 14 23 | 11

-1 | 11 14 | 23.

O/p :-	0	1	2	3	4	5	6	7	8
Sub1 :-	11				14				23
Sub2 :-		18				40			
Sub3 :-			19				29		
Sub4 :-				30				37	

Final Output :-
of Pass 1.

O/p :-	0	1	2	3	4	5	6	7	8
11	18	19	30	14	40	29	37	23	
11	18	19	30	14	40	29	37	23	

$$\text{gap} = \text{floor} \left(\frac{n}{2^2} \right) \text{ 2nd pass.}$$

(gap = 2).

Sub1 - $a[0], a[2], a[4], a[6], a[8]$

Sub2 - $a[1], a[3], a[5], a[7]$

0 1 2 3 4 5 6 7 8
 11 18 19 30 14 40 29 37 23
 11 19 14 29 23
 18 30 40 37

p & Pao-2

- 11 14 19 23 29.
 - 18 30 37 40

Final O/p of Pao-2

11, 18, 14, 30, 19, 37, 23, 40, 29.

2-2

$$\text{gap} = \text{floor}\left(\frac{n}{2^2}\right)$$

$$\text{gap} = 1.$$

$a[0]$ $a[1]$ $a[2]$ $a[3]$ $a[4]$ $a[5]$ $a[6]$ $a[7]$
 11 18 14 30 19 37 23 40 29

11 14 18 30 19 37 23 40 29
 11 14 18 30 19 37 23 40 29
 11 14 18 19 30 37 23 40 29
 11 14 18 19 23 30 37 40 29
 11 14 18 19 23 29 30 37 40
 11 14 18 19 23 29 30 37 40

Hashing (Address searching techniques)

Hash Table

It is a DS, which consists of an array in which data is accessed via a special index known as key. Hash function

It is used to store data in Hash Table, it establishes mapping between set of possible keys & position in hash table.

Outcome of Hash func is called Hash code / Hash value.

It'll be of integer type.

Applications of Hashing

Data Dictionary

Password verification on any postal

File stored on cloud (virtual storage).

Mapping b/w file name & file path Database system of any organization Cryptography Symbol table

All operations are performed with the help of # function.

A = 0.61 (99%)

Multiplication Method:

Operates in 2 steps:

(1) Multiplication method: \times the key by a constant A; where A is between 0 & 1. Then abstract subtract the fractional part of $A \times k$.

(2) Multiply extracted value by m & take float / integer of the result.

Example:

$$KA = K=107 \quad M=50, \\ A = 0.61$$

$$(1) \quad KA = 65.27$$

$$(2) \quad m \times 27 = 50 \times 0.27$$

$$\text{floor}(13.50) \rightarrow 13 \quad = 13.50$$

{ Must be always 6/0
0 to 50}.

Division Method:

Map the key (k) into one of the m slot by taking remainder of map key k into $m \bmod m$.
record

m - length of hash table.

0	72	89
1	51	37
2		
3		
4		
5	89	
6	51	
7		
8		
9		
10		
11		

Total size of m.

Certain values of m must be avoided
(power of 2 & power of 10)

Collision: 2 values (data) having same address.

$$60 \bmod 10 \\ 150 \bmod 10$$

Mid-Square Method:

- Square the key.
- Take mid part of the result as an address.

$$3111 \rightarrow \text{Key} \\ (3111)^2 = 967321$$

3111

Aage se gye toh Mid = 78
else. Mid = 83.
↳ (given here 78)

Digit-Folding Method:

Break the Key into small pieces
add them to get the hash
address.

Truncate from MSB or higher digit
based on size of table

$$h(\underline{\underline{823}}\underline{\underline{94564}}) = 823 + 94 + 784 = 1478,$$

$$h(\underline{\underline{871}}\underline{\underline{39465}}) = 871 + 39 + 465 = 1375$$

$$h(\underline{\underline{835}}\underline{\underline{67271}}) = 835 + 67 + 271 = 1173$$

Highest ASCII code. (127)

2 23, Page:

Now, if $m = 1000$.

i.e., size of table \neq address

$$\begin{aligned} & \therefore = 1478 \\ & = 375 \\ & = 173. \end{aligned}$$

For string:

suresh.

$m = 977$
(table size)

$$S + u + x + e + s + h, \\ \rightarrow 115 + 119 + 114 + 101 + 115 + 104 \\ m = 977$$

$$= 666.$$

$$= 666 \cdot m = 666 \cdot 977$$

= 84582.

If $m >$ data.
so; Multiply data by 11 ASCII code.

$$= 127 \times 666$$

$$= 84582 \cdot 9 \cdot m$$

$$= 84582,$$

[When more than 1 key, map to same # value in # table] collision

Resolution of Collision:

→ By separate chaining.

→ Open addressing

(1.) $5, 28, 19, 15, 20, 33, 12, 17, 10.$

$$h(k) = k \bmod m.$$

$$h(5) = 5 \bmod 9.$$

$$= 5$$

$$h(28) = 28 \bmod 9 = 1$$

$$h(19) = 19 \bmod 9 = 1$$

$$h(15) = 15 \bmod 9 = 6$$

$$h(20) = 20 \bmod 9 = 2$$

$$h(33) = 33 \bmod 9 = 6$$

$$h(12) = 12 \bmod 9 = 3$$

$$h(17) = 17 \bmod 9 = 8$$

$$h(10) = 10 \bmod 9 = 1$$

O/P of H.F - H Value

0	NULL				
1		10	-	19	-
2		-	20 X	-	
3		-	-	12 X	
4	NULL				
5		-	5 X	-	
6		-	-	15	-
7	NULL			-	
8		-	17 X	-	
9					

0 is advantage: Memory wasteage
[To store NULL address on link].

(2.) To search the addres when given (original # value) address is already occupied → Probing
(Always in → direction)

3 techniques of Probing:

- Linear
- Quadratic
- Double Hashing

Linear:

$$h(k, i) = [h'(k) + i] \bmod m.$$

$$h'(k) = k \bmod m \quad \text{for } i=0, 1, 2, \dots, m-1$$

The given keys are:

26, 37, 59, 76, 65, 86,

$$m = 11.$$

$$h'(26) = 26 \bmod 11 = 4$$

$$h'(37) = 37 \bmod 11 = 4$$

$$h'(59) = 59 \bmod 11 = 4.$$

$$h'(76) = 76 \bmod 11 = 10.$$

$$h'(65) = 65 \bmod 11 = 10.$$

$$h'(86) = 86 \bmod 11 = 9$$

$$h(k, i) = h(26, 0).$$

$$= [4 + 0] \bmod 11$$

$$= 4 \bmod 11$$

$\Rightarrow T(4)$.

Date: _____ Page: _____

Date: _____ Page: _____

$$h(37, 0) = (4+0) \bmod 11$$

$$h(59, 0) = (4 \Rightarrow T(4)) \text{ occupied}$$

$$\begin{aligned} h(37, 1) &= (4+1) \bmod 11 \\ &= 5 \bmod 11 \\ &\Rightarrow T(5). \end{aligned}$$

$$\begin{aligned} h(59, 1) &= 5 \bmod 11 \\ &\Rightarrow T(5) \text{ occupied} \end{aligned}$$

$$\begin{aligned} h(59, 2) &= T(6). \\ h(76, 0) &= (10+0) \bmod 11 \\ &= 10 \bmod 11 \\ &\Rightarrow T(10). \end{aligned}$$

$$\begin{aligned} h(65, 0) &= 10 \bmod 11 \\ &\Rightarrow T(10) \text{ occupied} \end{aligned}$$

$$\begin{aligned} h(65, 1) &= 11 \bmod 11 \\ &\Rightarrow T(11). \end{aligned}$$

$$\begin{aligned} h(86, 0) &= 9 \bmod 11 \\ &\Rightarrow T(9). \end{aligned}$$

0	65
1	NULL
2	NULL
3	NULL
4	26
5	37
6	59
7	NULL
8	NULL
9	86
10	76

D disadvantage
— Searching tym of index (7).

To overcome this linear; we'll use

(1) Quadratic \circlearrowleft

$$h(k, i) = [h'(k) + c_1 i + c_2 i^2] \text{ mod } m.$$

where c_1 & c_2 are constant.

$$h'(k) = k \text{ mod } m \text{ for } i=0, 1, 2, \dots, m-1.$$

linear $\vdash h, h+1, h+2, h+3, \dots$

quadratic $\vdash h, h+1, h+4, h+9, \dots$

Double Hashing \downarrow

Why D# is best?

Bcz quad. thi kahi na kahi linearly work krti h.

2 # functions are used there.