

1) Two applications of Sorting and data Structures.

Ans Sorting is the process of arranging data in a specific order (ascending or descending).

- Database Query Optimization → Sorting allows efficient searching and merging.

Ex → Sorting student records by roll no helps in faster binary search.

- Data Analysis → Sorting is used to find medians, rankings and duplicate detection.

Ex → In a e-commerce site, products are sorted by price or rating.

2) Static Hashing vs Dynamic Hashing

- Ans
- Static Hashing → The hash table size is fixed. If records increase, collision arises.

Ex → Student ID stored in a fixed table.

- Dynamic Hashing → The hash table expands/shrinks dynamically depending on data growth.

Ex → Extendible hashing in databases.

3) Internal vs External Sorting

- Ans
- Internal Sorting → data is stored entirely in main memory (RAM). Suitable for small datasets.

Ex → Quicksort on an array of 1000 elements.

- External Sorting → used when data is too large to fit in main memory.

fit in memory, Sorting is done using both memory and disk.

Ex → External merge sort for sorting $\log B \log n$

4) Buffer management in external sort

Ans • In External Sorting, disk access is expensive. Buffer management allocates blocks of memory as I/O buffers. Instead of reading/writing one record at a time large chunks are moved reducing I/O operations.

5) External Sorting with example

Ans. External Sorting is used when data is larger than available memory.

Steps

- 1) Divide large file into chunks (runs).
- 2) Sort each run in memory.
- 3) merge runs until final sorted file is obtained.

Ex → Storing 1 TB log file in Binary

6) Advantage of Double hashing over Linear Probing

Ans • Linear Probing: Collisions are resolved by searching next slots sequentially, which causes primary clustering.

• Double hashing uses a second hash function to calculate probe step size, reducing clustering and spreading keys uniformly.

Ex → $H_1(k) = k \bmod 7$, $H_2(k) = 1 + (k \bmod 5)$.

7) Folding method

Ans Table size = 10.

- Key 987654321 $\rightarrow 987 + 654 + 321 = 1962 \text{ mod } 10 = 2$.
- Key 5643231 $\rightarrow 564 + 323 + 1 = 888 \rightarrow 888 \text{ mod } 10 = 8$.
- Key 3478654 $\rightarrow 347 + 865 + 4 = 1216 \rightarrow 1216 \text{ mod } 10 = 6$.

8) Double Hashing.

Ans $H_1(k) = k \text{ mod } 13$, $H_2(k) = 12 - (k \text{ mod } 12)$
 $H(k) = (H_1(k) + i * H_2(k)) \text{ mod } 13$

44	41			18	32	59	73	22	31		
0	1	2	3	4	5	6	7	8	9	10	11

$$18 \text{ mod } 13 = 5$$

$$41 \text{ mod } 13 = 2$$

$$22 \text{ mod } 13 = 9$$

$$44 \text{ mod } 13 = 5 \rightarrow$$

$$59 \text{ mod } 13 = 7$$

$$32 \text{ mod } 13 = 6$$

$$31 \text{ mod } 13 = 5 \rightarrow$$

$$H_2 = 12 - (31 \text{ mod } 12) = 12 - 7 = 5$$

$$i=1 \rightarrow (5+5) \text{ mod } 13 = 10$$

$$73 \text{ mod } 13 = 8$$

9) Open Addressing Example.

Ans $H(k) = k \text{ mod } 6$, $k = 24, 19, 32, 44, 57, 5$

$$24 \text{ mod } 6 = 0$$

1	1	1	2			
24	19	32	44	57		
0	1	2	3	4	5	

$$24 \bmod 6 = 0$$

$$19 \bmod 6 = 1$$

$$32 \bmod 6 = 2$$

$$44 \bmod 6 = 2 \rightarrow \text{next free} = 3$$

$$57 \bmod 6 = 3 \rightarrow \text{next free} = 4$$

10) Mid-Square Method

Ans $k = 625$, Table size = 100

$$625^2 = 390625$$

$$\text{mid-digits} = 06$$

$$06 \bmod 100 = 6$$

$$\text{Hash value} = 6$$

11) k-way merge ($k=3$)

Ans $k=3$ means merging 3 runs simultaneously.

Ex -)

Run 1 = [], 7, 10], Run 2 = [1, 4, 8], Run 3 = [3, 5, 9]

Process: Pick smallest \rightarrow [1, 2, 3, 4, 5, 7, 8, 9, 10].

~~like Run 1~~

12) Linear probing - Example

Ans Table size = 7, $h(k) = k \bmod 7$

$k = 50, 700, 76, 85, 92$

$$50 \bmod 7 = 1$$

$$700 \bmod 7 = 0$$

$$76 \bmod 7 = 6$$

$$85 \bmod 7 = 1 \rightarrow \text{collision} = 2$$

1	2	3	4	5	6	7
700	50	85	92			76
0	1	2	3	4	5	6

$$92 \bmod 7 = 1 \rightarrow \text{collision} = 3$$

13) Comparison of hashing techniques.

Ans: i) Mid-Square \rightarrow Spreads keys well but Depends on mid numbers

ii) folding \rightarrow Breaks keys into parts and adds, good for long no

iii) Data Analysis \rightarrow Selects specific digits, useful if distribution is uniform.

iv) Double Hashing \rightarrow Best for reducing clustering, gives uniform distribution.

• •

14) External merge sort example.

Ans Run 1 = [3, 7, 10]

Run 2 = [1, 4, 8]

Run 3 = [2, 5, 9]

Smallest \rightarrow [1, 2, 3] =

next small \rightarrow [4, 5, 7]

next = [8, 9, 10]

k = [1, 2, 3, 4, 5, 7, 8, 9, 10].

15) Time complexity of external sort.

Ans formula = $O(N \log_k (N/M))$, where

$N = 7000$, $M = 100$, block = 100 $\rightarrow k = 700/100 = 7$

No. runs = $N / \text{block} = 7000 / 100 = 70$

passes = $\log_7 (70) \approx 2$.

Time complexity $\approx O(7000 \times 2) = O(14,000)$.