Exam
Make - up
April 16 - 18, 2018

Name: RITIKA KUMARI

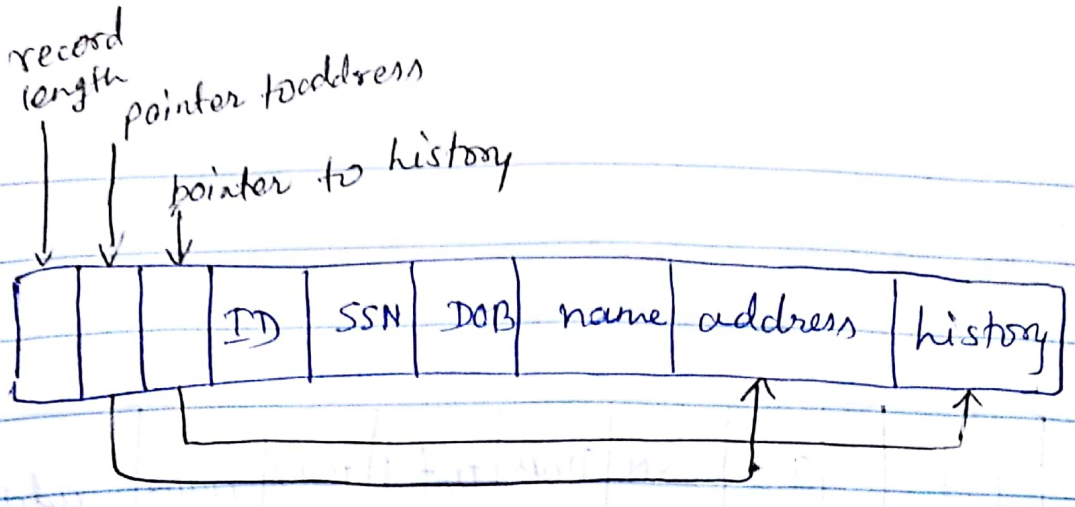
CWID – A20414073

- I affirm my awareness of the Standards of the Illinois Institute of Technology Honor Code

Sign: Ritika kumari

Q 1.1

Soln



record length = 4 bytes

DOB = 10 bytes

SSN = 10 bytes

Patient ID = 10 bytes

pointer to address = 4 bytes

" " patient history = 4 bytes

∴ Total no. of bytes needed = 4 + 10 + 10 + 10 + 4 + 4

= 12 + 30 = 42 bytes Ans

Q 1.2.

Soln:

Avg. bytes for name = $\frac{10+50}{2} = \frac{60}{2} = 30$ bytes

Avg. bytes for address = $\frac{20+80}{2} = \frac{100}{2} = 50$ bytes

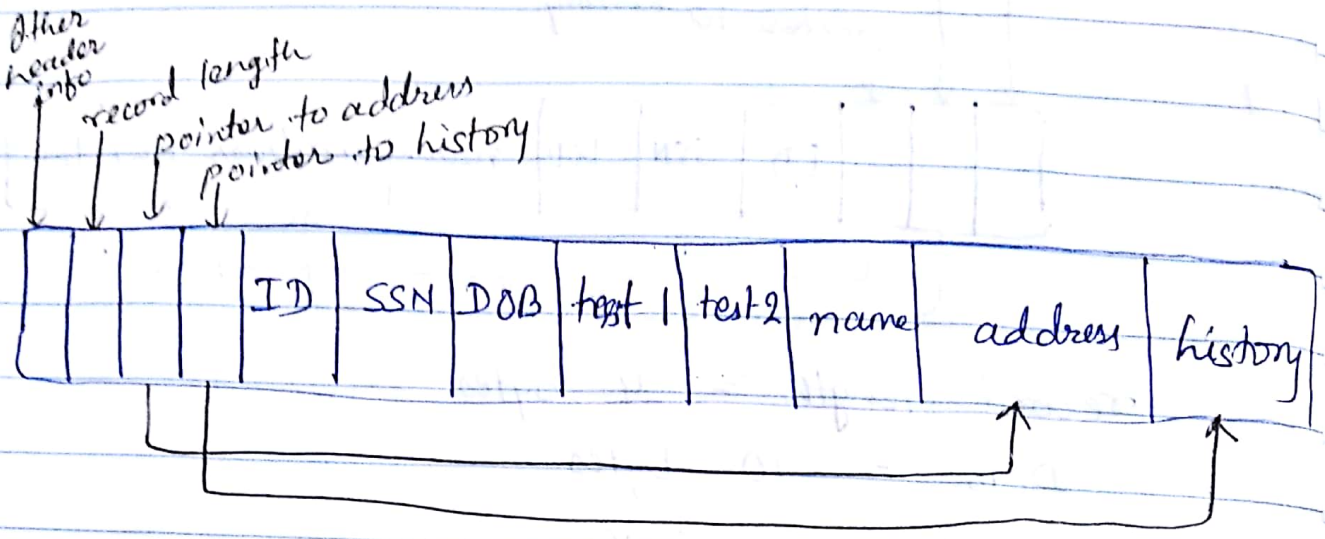
Avg bytes for history = $\frac{0+1000}{2} = 500$ bytes

∴ Average length of record

= 4 + 10 + 10 + 10 + 4 + 4 + 30 + 50 + 500

= 622 bytes Ans

Q 1.3 (a)



layout of record :

record length = 4 bytes

pointer to address = 4 bytes

" " patient history = 4 bytes

DOB = 10 bytes

SSN = 10 bytes

Patient ID = 10 bytes

name = variable length

address = variable length

Patient history = variable length

Cholesterol test Date = 16 bytes

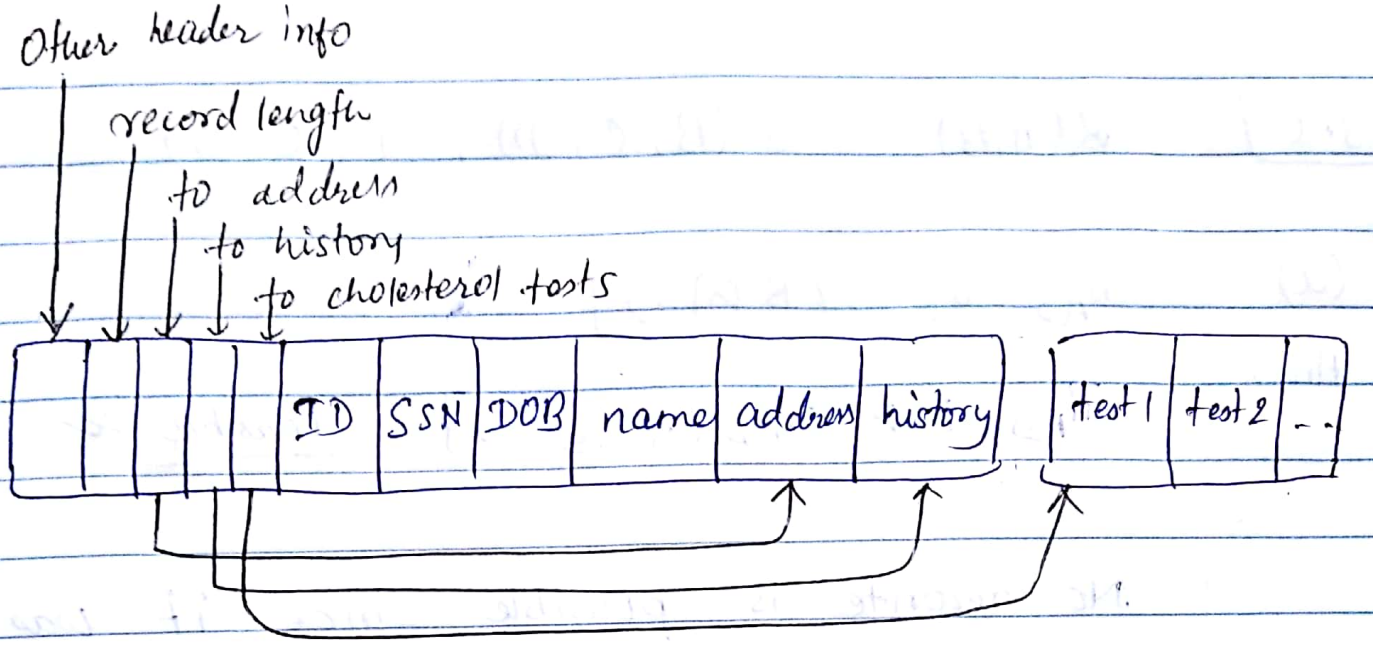
Cholesterol Test Result = 4 bytes

Cholesterol Test Date = 16 bytes

Cholesterol Test Result = 4 bytes

....

13(b)



layout of a record:

record length = 4 bytes

pointer to address = 4 bytes

pointer to patient History = 4 bytes

DOB = 10 bytes

patient ID = 10 bytes

SSN = 10 bytes

name = variable length

address = variable length

patient history = variable length

pointer to Cholesterol tests = 4 bytes

Since Cholesterol Tests is stored on a separate block with layout:

Cholesterol Test Date = 16 bytes

Cholesterol Test Result = 4 bytes

Cholesterol Test Date = 16 byts

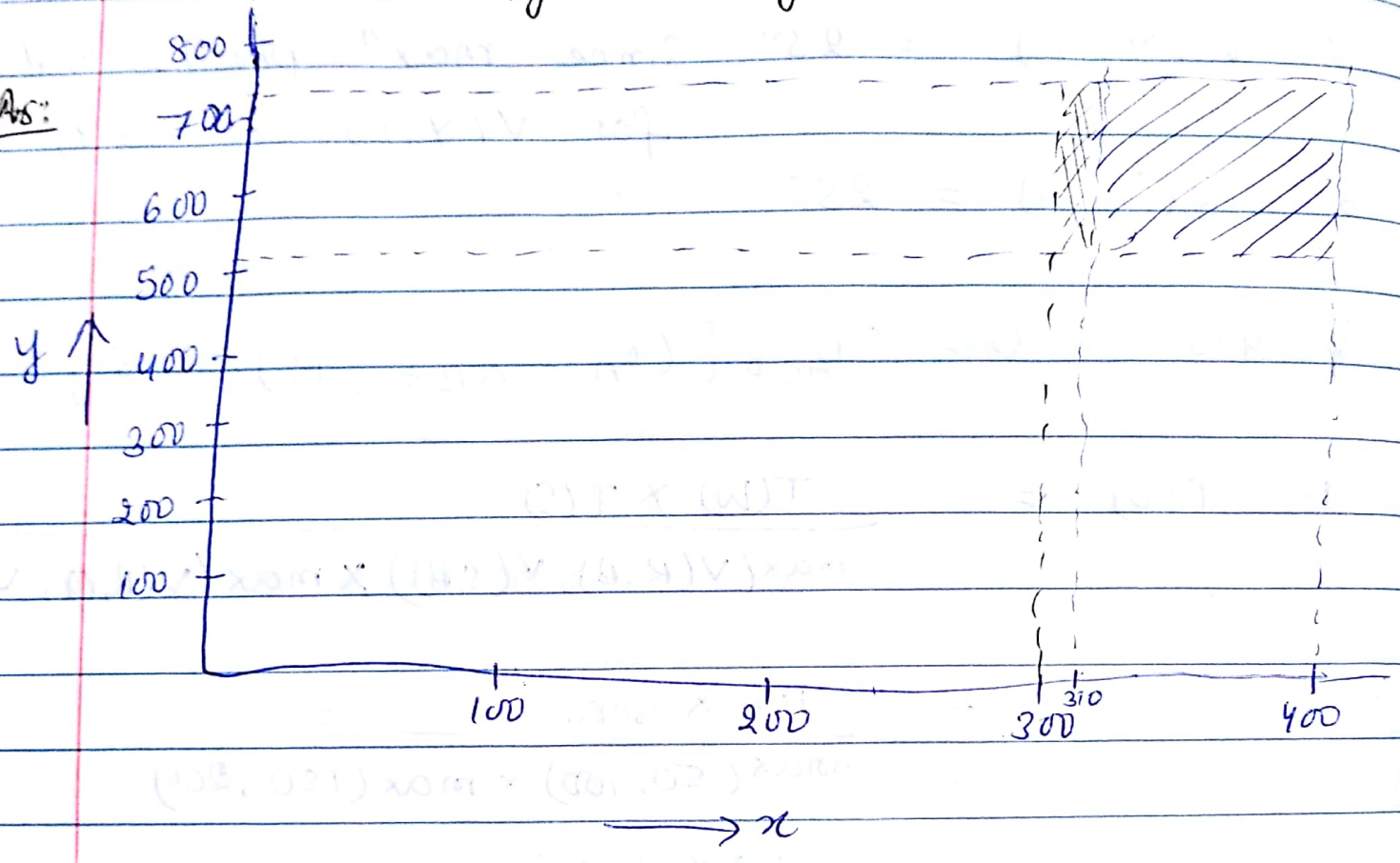
Cholesterol Test Result = 4 bytes

....

Q 2.1

(a) select * from R where $310 < x$ and $x < 400$ and $520 < y$ and $y < 730$;

Ans:



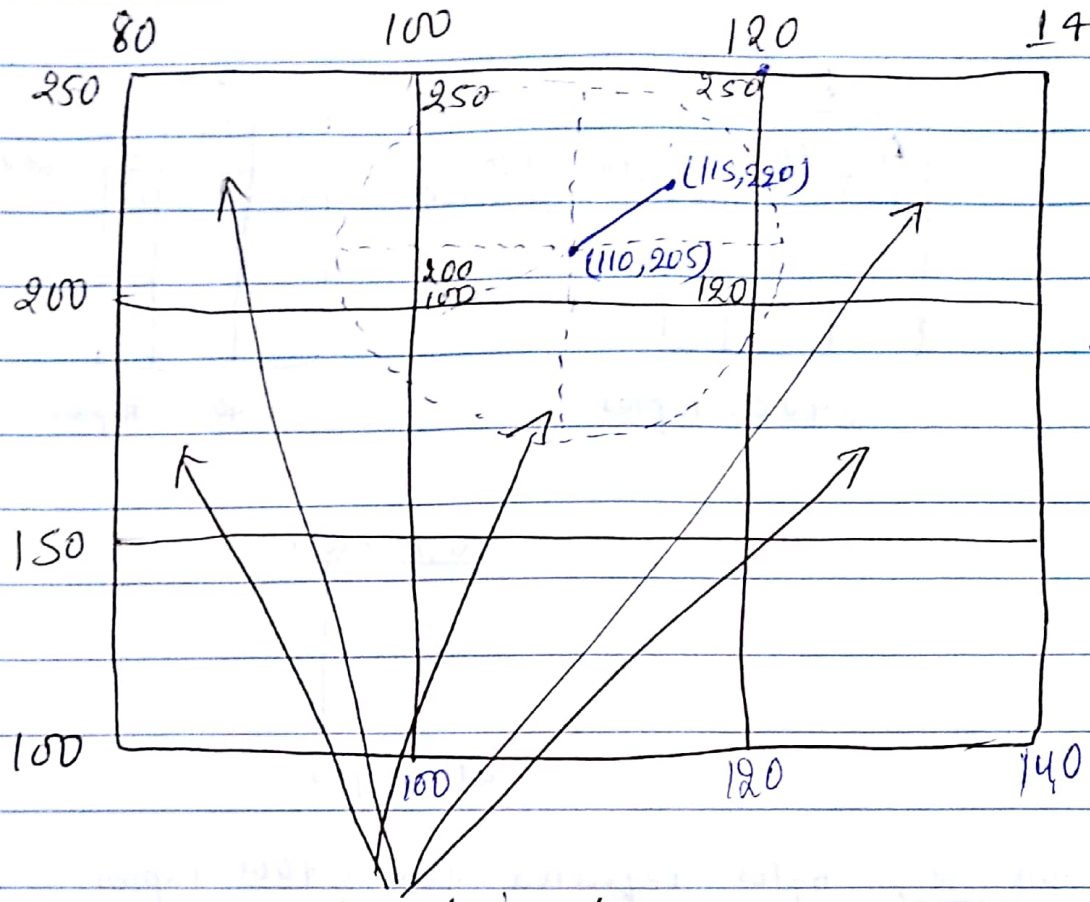
x : [300-320] [320-340] [340-360] [360-380] [380-400]

y : [500-550] [550-600] [600-650] [650-700] [700-750]

∴ Total no. of buckets = 5 × 5 = 25 Ans

2.1

(b)



$r = \sqrt{225+25}$

$= \sqrt{250}$

$= 16$

∴ The buckets are :

[80-100] [150-200]

[80-100] [200-250]

[100-120] [150-200]

[120-140] [150-200]

[120-140] [200-250] Ans

3.1 $R(A,B)$, $S(B,C,D)$, $T(C,D)$

(a) $\Pi_{AD}[\sigma_{C=5}(R \bowtie S)]$

Ans: $\Pi_{AD}[R \bowtie \Pi_{BD}(\sigma_{C=5} S)]$ Feasible ✓

No rewrite is possible since it was not specified whether the selection could also be pushed down.

(b) $[R \bowtie (\sigma_{C=2} S)] \cup [(\sigma_{A=1} R) \bowtie S)]$

Ans: Yes it is feasible and we can rewrite the expression as:

$\sigma_{(C=2 \vee A=1)} [R \bowtie S]$

(c) $\sigma_{D=2}[(\sigma_{C=2} S) \bowtie (\sigma_{C=1} T)]$

Ans: Not feasible because if we select all rows with $C=2$ from S and all rows with $C=1$ from T, then we will not get any output in natural join on C and D. We will avoid running the entire query in this case.

Part 4: $R(A, B, C)$ $S(B, C, D)$

$$U = \Pi_{ACD}\left[\left(\sigma_{A=3 \wedge B=5} R\right) \bowtie S\right]$$

$T(R) = 100000$; $V(R, A) = 20$; $V(R, B) = 50$, $V(R, C) = 150$

$T(S) = 5000$; $V(S, B) = 100$; $V(S, C) = 200$; $V(S, D) = 30$

each attribute = 10 byte in size

Q. 4.1. $W = \sigma_{A=3 \wedge B=5} R$.

1. $T(W) = \frac{1}{V(R, A)} \times \frac{1}{V(R, B)} \times T(R)$

$= \frac{1}{20} \times \frac{1}{50} \times 100000$

$= 100$ Ans.

2. $S(W)$

R has three attributes A, B & C in W

$\therefore S(W) = 10 + 10 + 10$

$= 30$ bytes

3. $V(W, A)$

Ans: 1

4. $V(W, B)$

Ans: 1

5. $V(W, C)$

Ans. 100 Since the max^m possible value for $V(W, C)$ is $T(W)$.

Q. 42. $Y = W \bowtie S$

1. $$T(Y) = \frac{T(W) \times T(S)}{\max(V(R,B), V(S,B)) \times \max(V(R,C), V(S,C))}$$

$$= \frac{100 \times 5000}{\max(50, 100) \times \max(150, 200)}$$

$$= \frac{\cancel{100} \times 5000}{\cancel{100} \times 200} = 25 \text{ Ans}$$

2. $S(Y) = S(A) + S(B) + S(C) + S(D)$

$= 10 + 10 + 10 + 10$

$= 40$ bytes Ans

3. $V(Y, A) = 1$

4 $V(Y, B) = 1$

5. $V(Y, C) = 25$ Since max^m possible value for $V(Y, C)$ is $T(Y)$.

6. $V(Y, D) = 25$ " " " "

Q. 4.3. $U = \Pi_{ACD}[(\sigma_{A=3 \wedge B=5} R) \bowtie S]$

1. $T(U) = \dfrac{T(W) \times T(S)}{\max(V(R,B), V(S,B)) \times \max(V(R,C), V(S,C))}$

$= \dfrac{100 \times 5000}{\max(50, 100) \times \max(150, 200)}$

$= \dfrac{\cancel{100} \times 50\cancel{0}\cancel{0}}{\cancel{100} \times 2\cancel{0}\cancel{0}} = 25$ Ans

2. $S(U)$

$S(U) = S(A) + S(C) + S(D)$

$= 10 + 10 + 10 = 30$ Ans

3. $V(U, A) = 1$ Ans

4 $V(U, C) = 25$ Since max^m possible value for $V(U, C)$ is $T(U)$.

5. $V(U, D) = 25$ " " "

Q 5.1:

Soln:

R

A	B	C
10	10	10

Total = 30 bytes each record

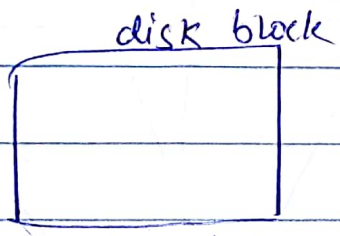
64000 tuples

S

A	D	E
10	15	15

Total = 40 bytes each record

64000 tuples



6400 bytes

For R, bytes required for 64000 tuples = 30 × 64000

= 30 × 64000 bytes

∴ Blocks required for 64000 tuples

$$B(R) = \frac{30 \times 64000}{6400} = 300 \text{ blocks}$$

For S, bytes required for 64000 tuples = 40 × 64000

∴ Blocks required for 64000 tuples

$$B(S) = \frac{40 \times 64000}{6400} = 400 \text{ blocks}$$

Since R and S are not sorted, we will go with two pass Hash algorithm.

∴ No. of IOs required = 3(B(R) + B(S))

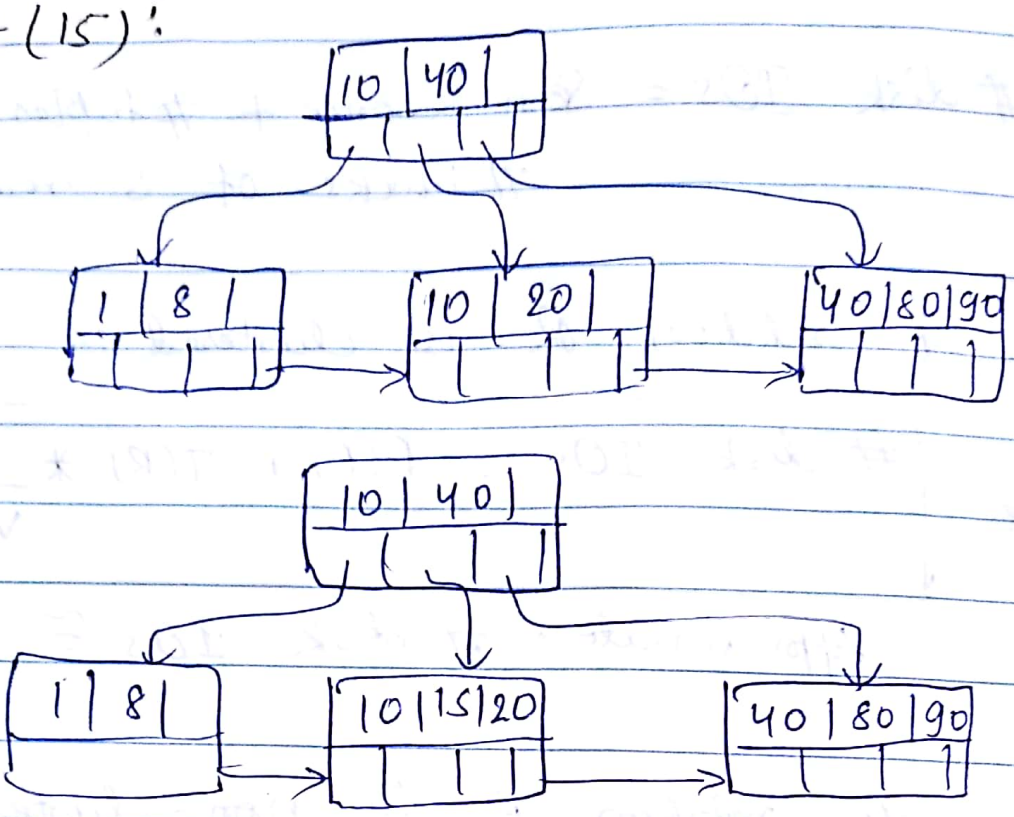
= 3(300 + 400)

= 3 X 700

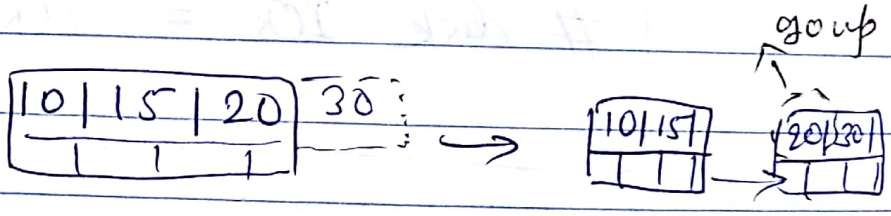
= 2100 Ans

Q 6.1.

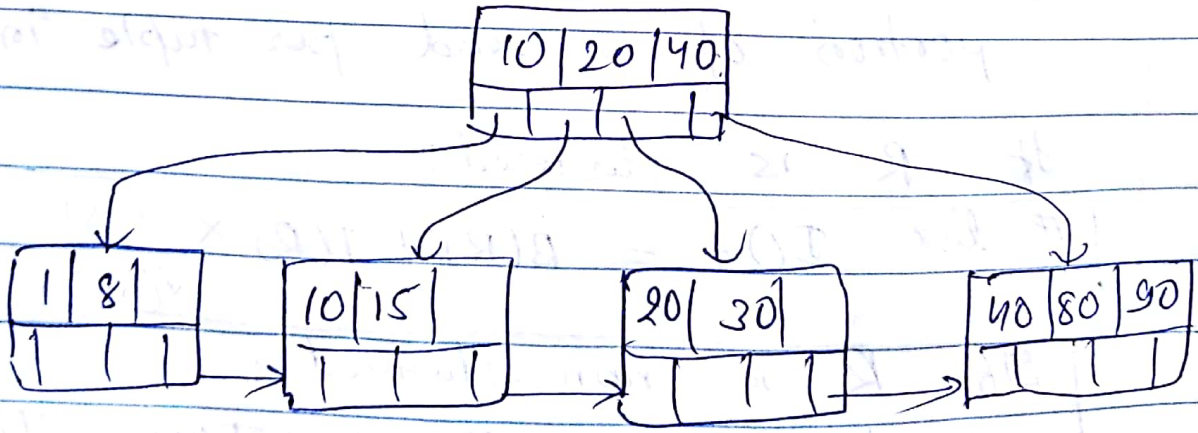
insert (15):



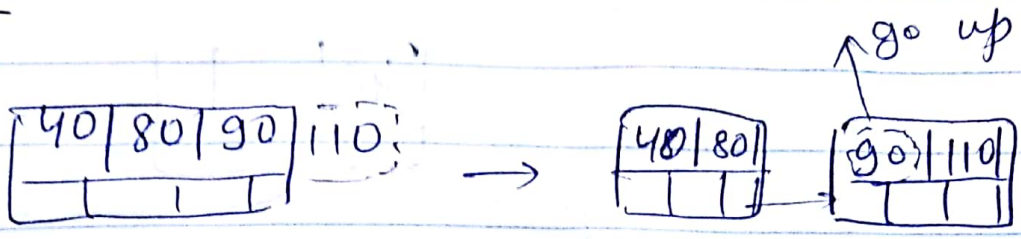
insert (30):



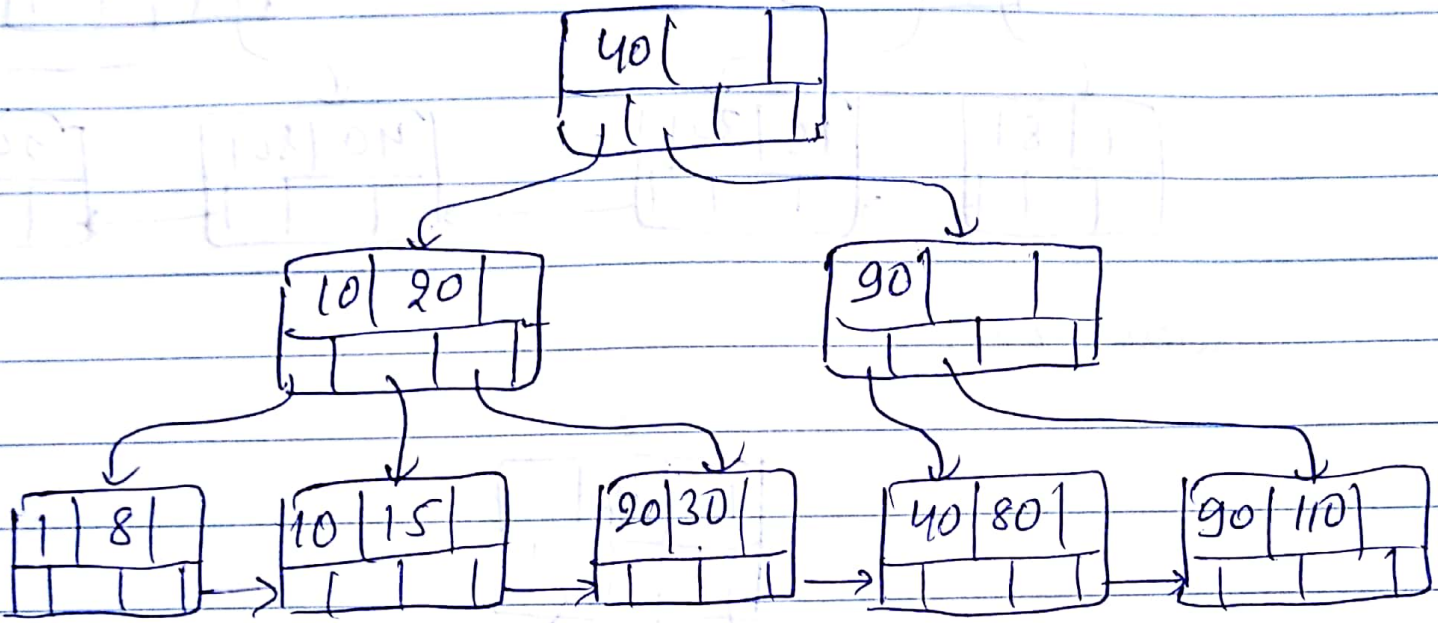
∴ final tree after inserting 30



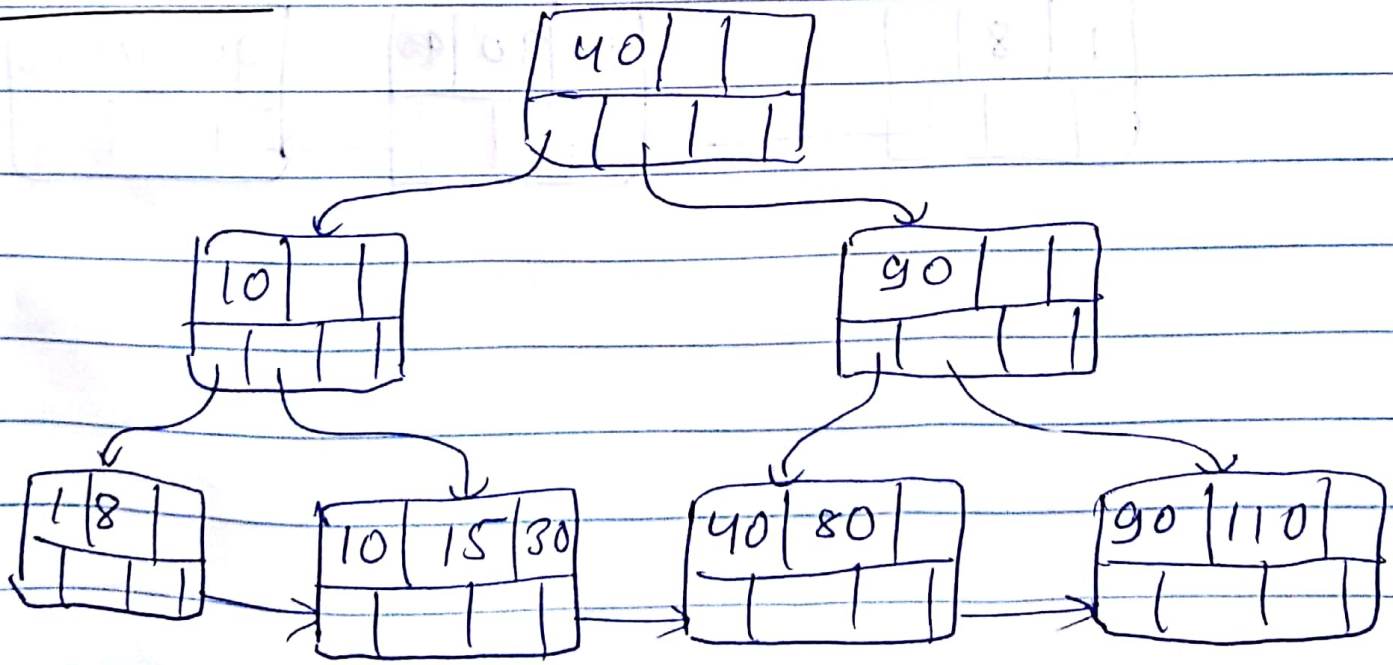
insert (110) —



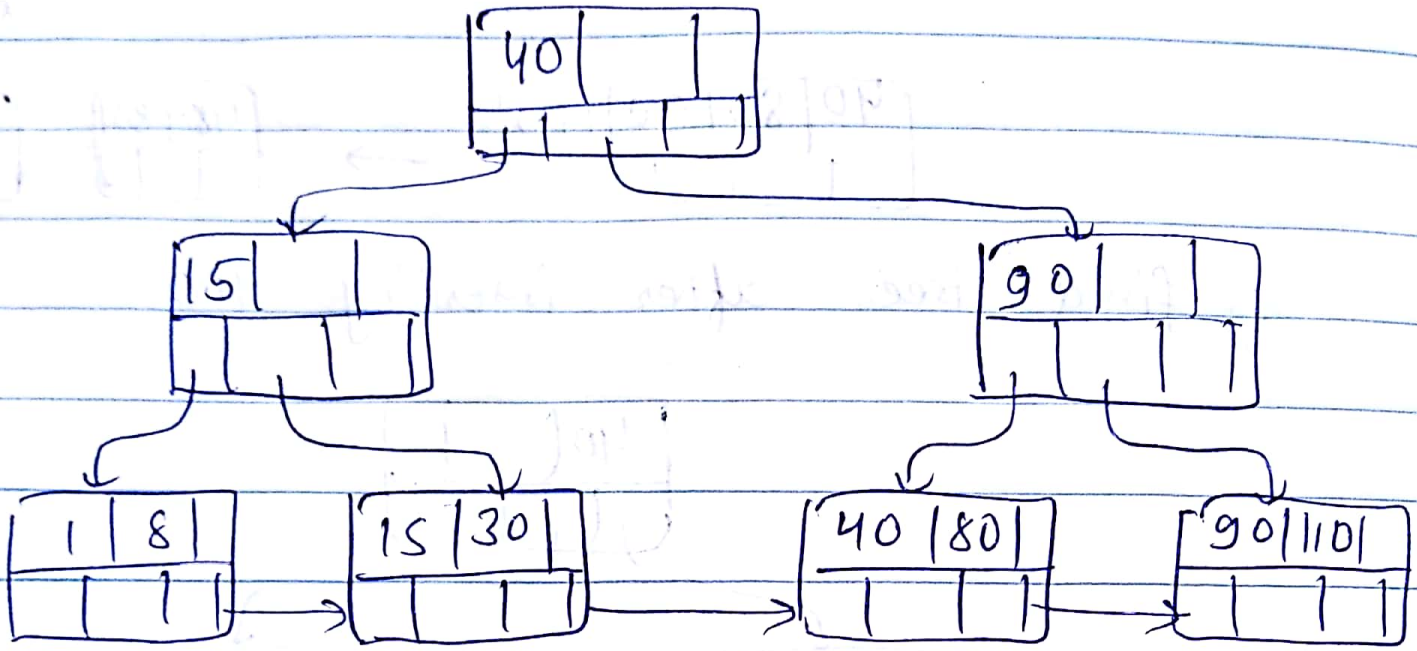
∴ final tree after inserting 110



delete (20)



delete (10)



delete (80):