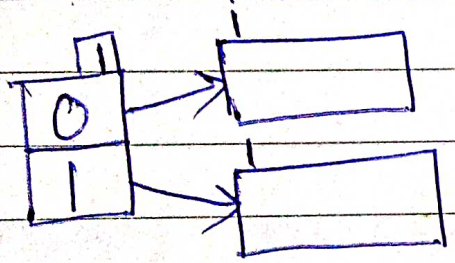
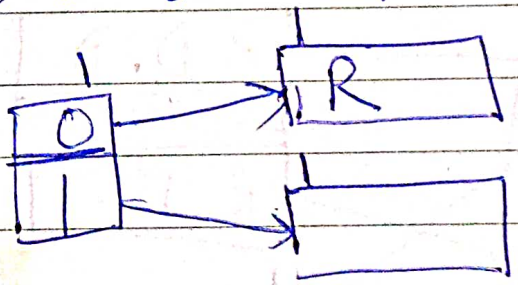
1a) R → 82 → 1010010
P → 80 → 1010000
D → 68 → 1000100
J → 74 → 1001010
e → 67 → 1000011
G → 71 → 1000111
D → 68 → 1000100
A → 65 → 1000001

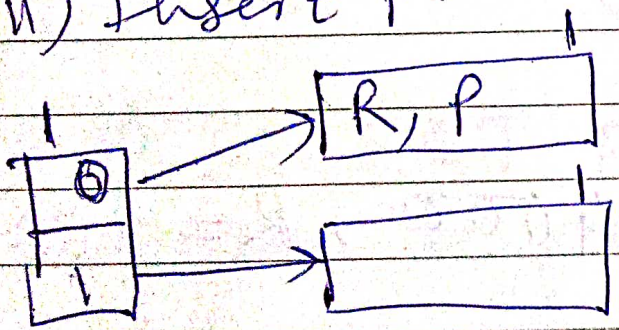
Bucket Size = 3.
LD = GD = 1.



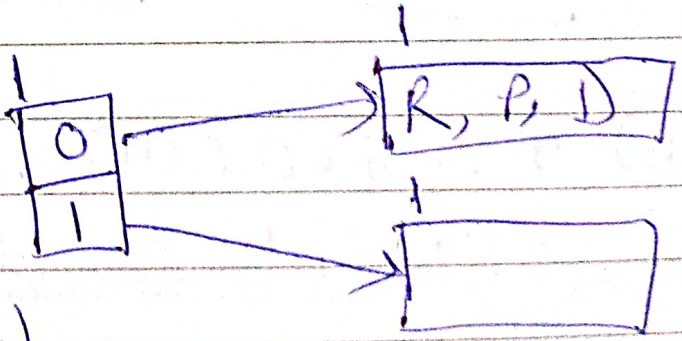
i) Insert R



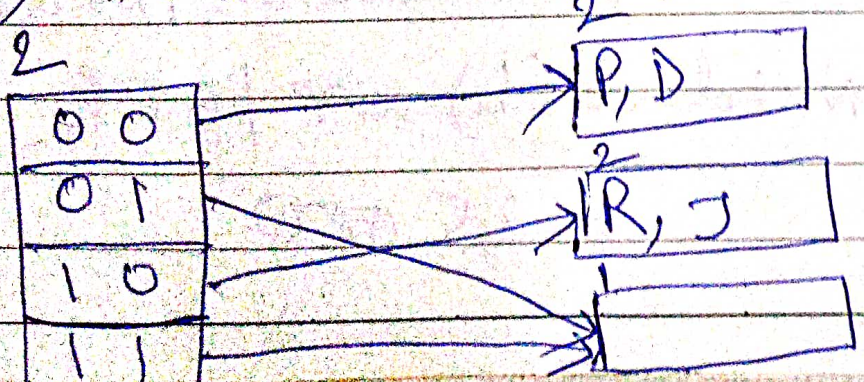
ii) Insert P.



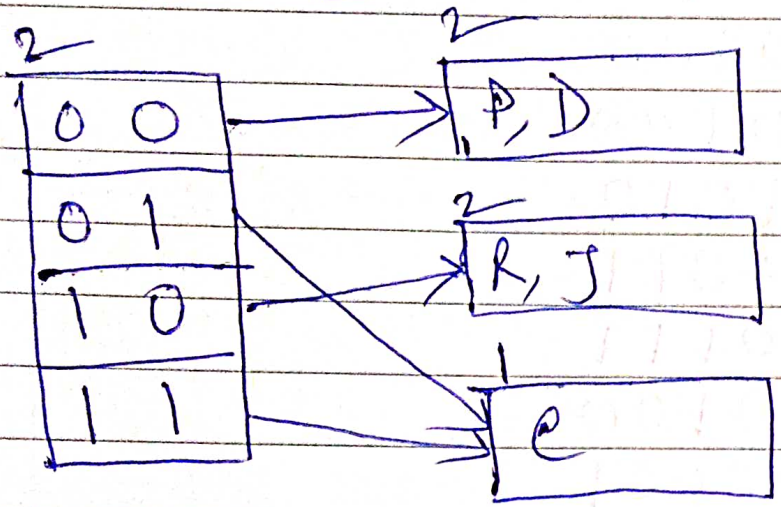
iii) Insert D.



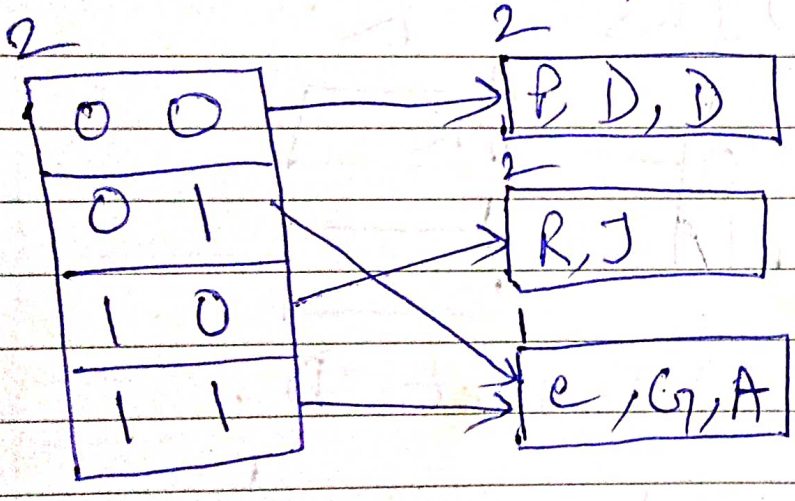
iv) Insert J (Overflow).



v) Insert C



vi) Insert G, D, A



b) $r = 500000$, $B = 4000$, $R = 200$

Without Index

$BFR = B/R = 20$ records/Block

No. of blocks Needed $(b) = r/BFR$

$= 25000$

A binary search on the data file
$= \log_2 b = \log_2 25000 = (15)$

Using Clustering Index

Index entry = 1500
Index size = 5 + 7 = 12 bytes.
BFR = 4000/12 = 334
No. of blocks needed (b) = 1500/334
= 5

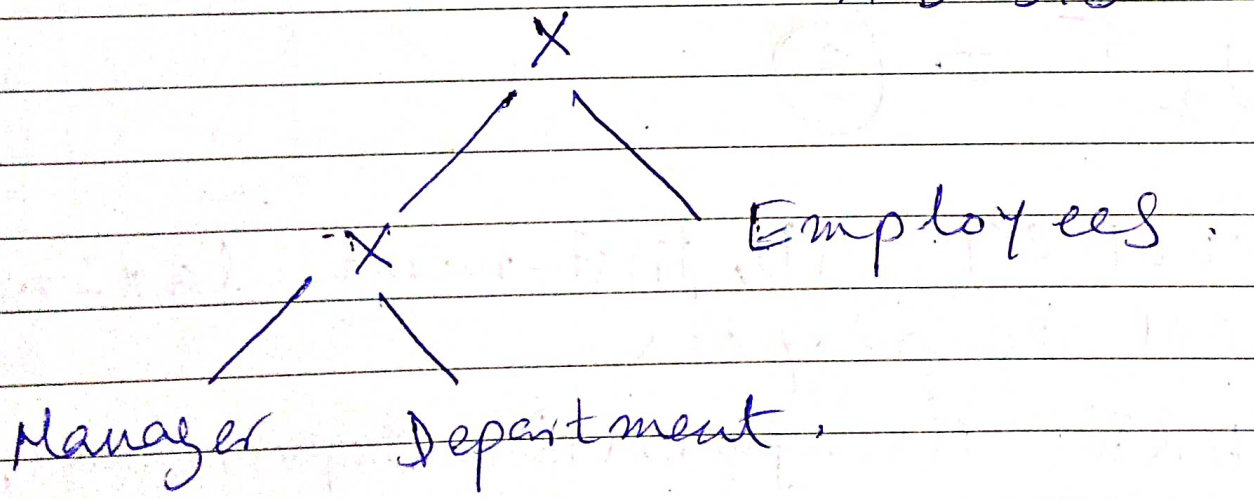
Binary search on ~~data~~ Index file.
$\log_2 5 = (3)$

2a) SELECT EID, first-name, last-name
FROM employees
WHERE Manager.MID = Department.MI
AND Department.DID = Employees.DID
AND Manager-name = 'Sagar'.

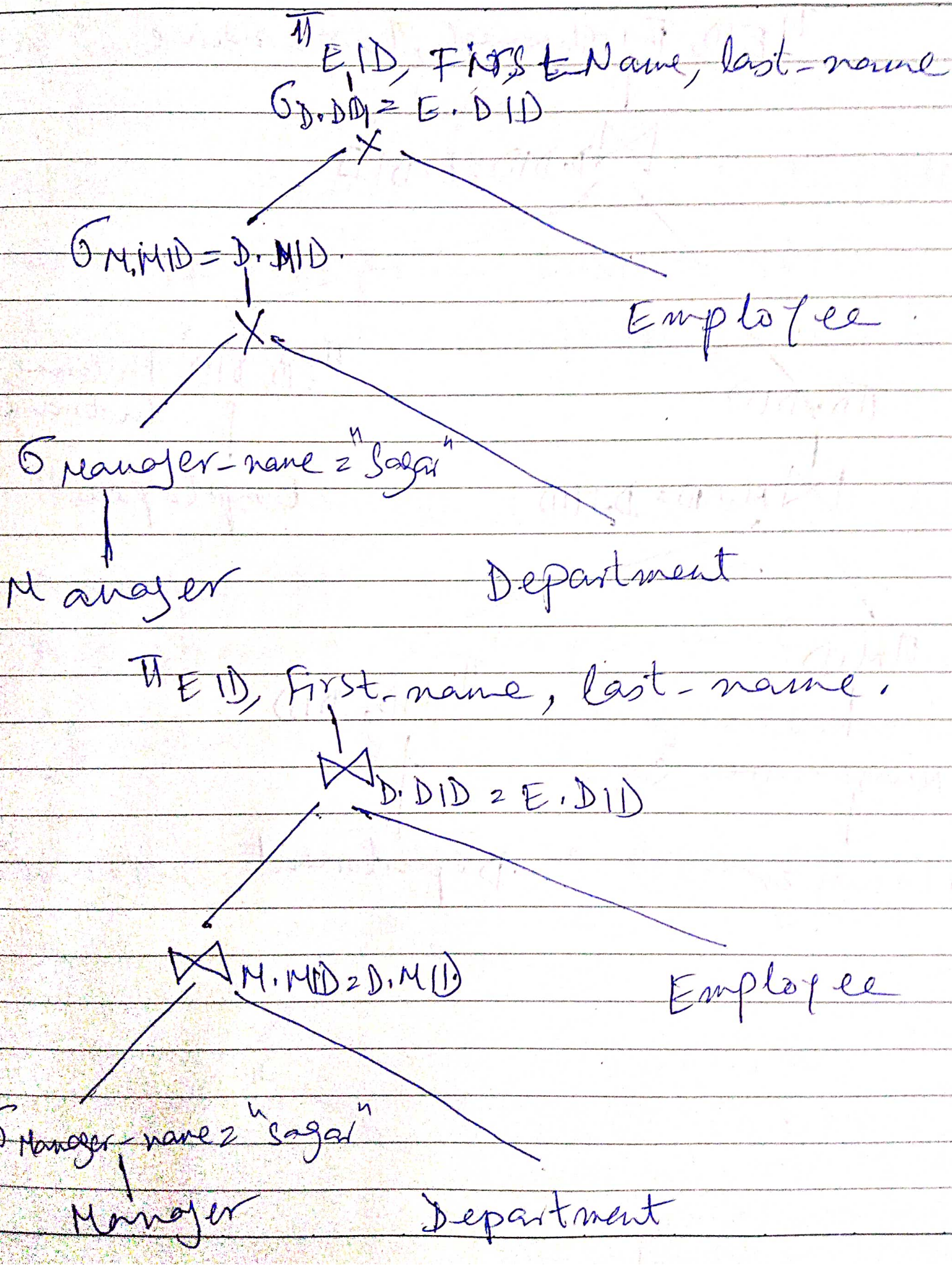
Π EID, first_name, last_name

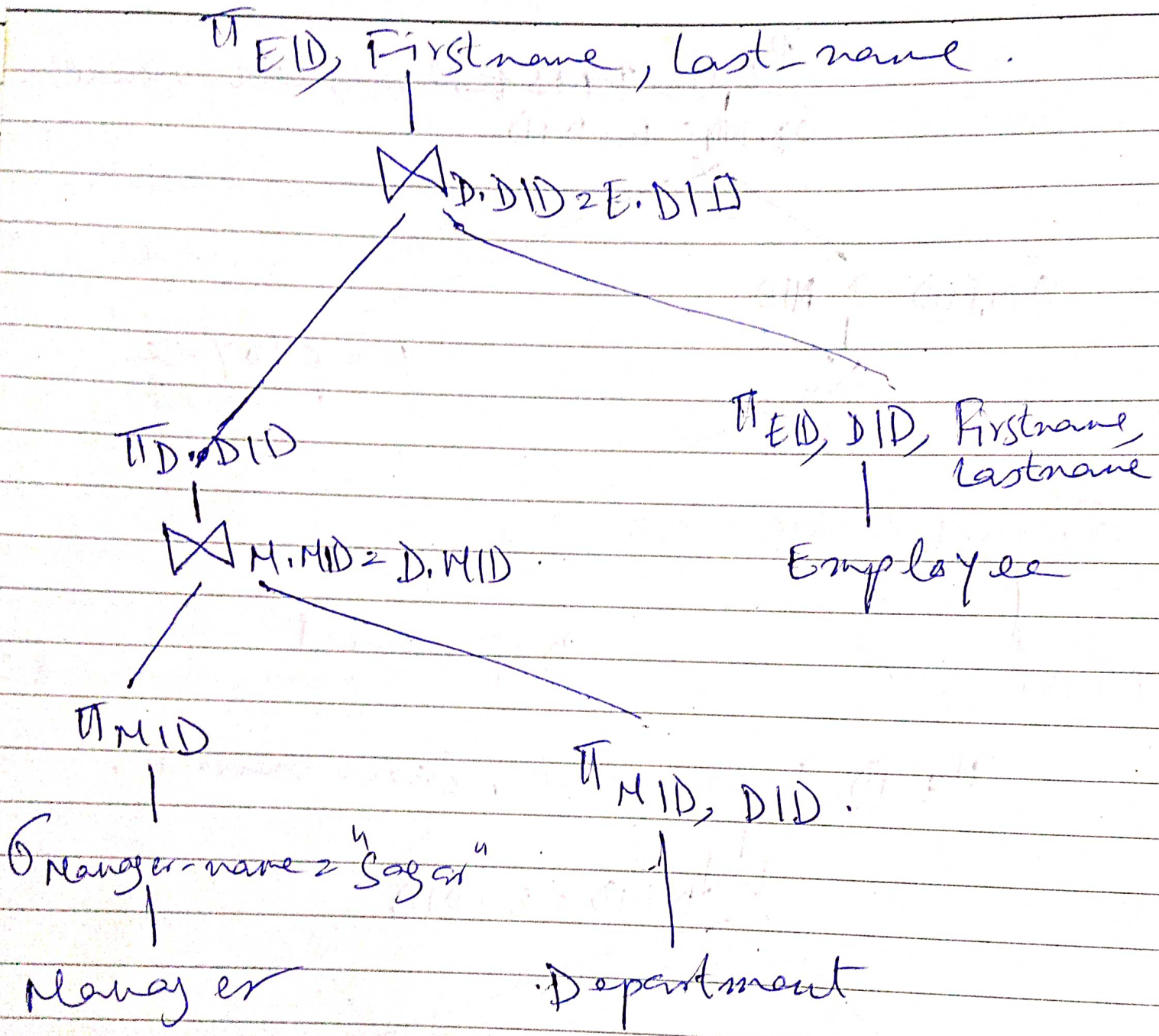
σ Manager_name = "Sagar" $\wedge$
Manager.MID = Department.MID $\wedge$
Department.DID = Employees.DID.
((Manager $\times$ Department) $\times$ Employees)

Π EID, first_name, last_name.
|
σ Manager_name = "Sagar" $\wedge$ M.MID = D.MID
$\wedge$ D.DID = E.DID.



[M = Manager,
D = Department
E = Employee]


Π E.ID, FIRST Name, last-name.
σ D.DID = E.DID
X
σ M.MID = D.MID
X
Employee
σ manager-name = "Sagar"
Manager
Department
Π E.ID, First-name, last-name.
⋈ D.DID = E.DID
⋈ M.MID = D.MID
Employee
σ Manager-name = "Sagar"
Manager
Department


Π EID, Firstname, Last_name.
⋈ D.DID = E.DID
Π D.DID
Π EID, DID, Firstname, Lastname.
Employee
⋈ M.MID = D.MID
Π MID
Π MID, DID.
σ Manager_name = "Sagar"
Manager
Department

No. 7

2b) SELECE EID, EName
FROM Employee
WHERE Product. Warehouse_location = "Delhi"
AND. product. PID = customer. PID
AND customer. CID = Employee. CID.

ii) $\Pi_{EID, EName}(\sigma_{Warehouse\_location = "Delhi" \wedge Product.PID = customer.PID \wedge customer.CID = Employee.CID}(Employee \times (Product \times customer)))$

$\Pi_{EID, EName}$

$\sigma_{Warehouse\_location = "Delhi" \wedge P.PID = c.PID \wedge c.CID = E.CID}$

$\times$

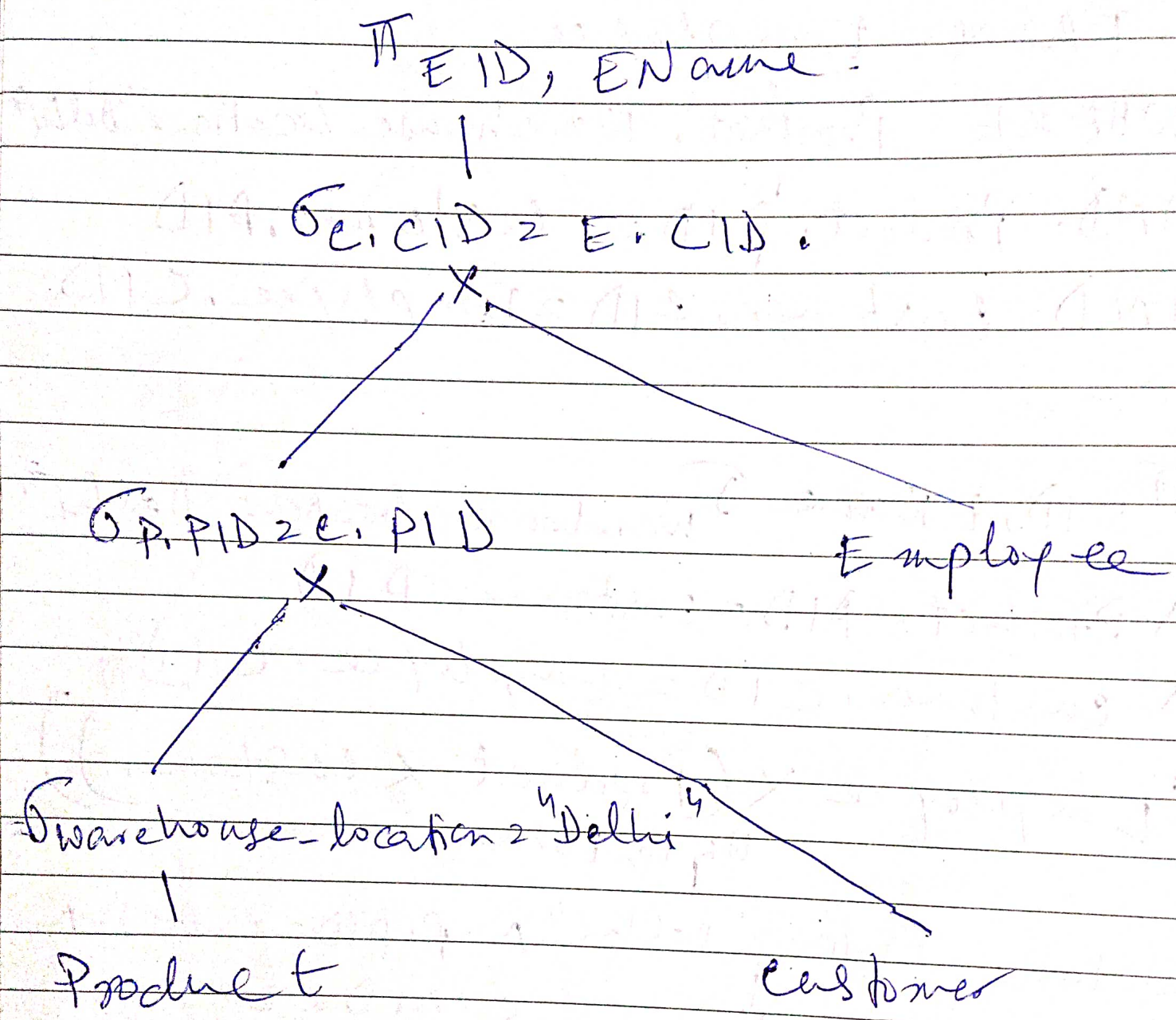
Employee

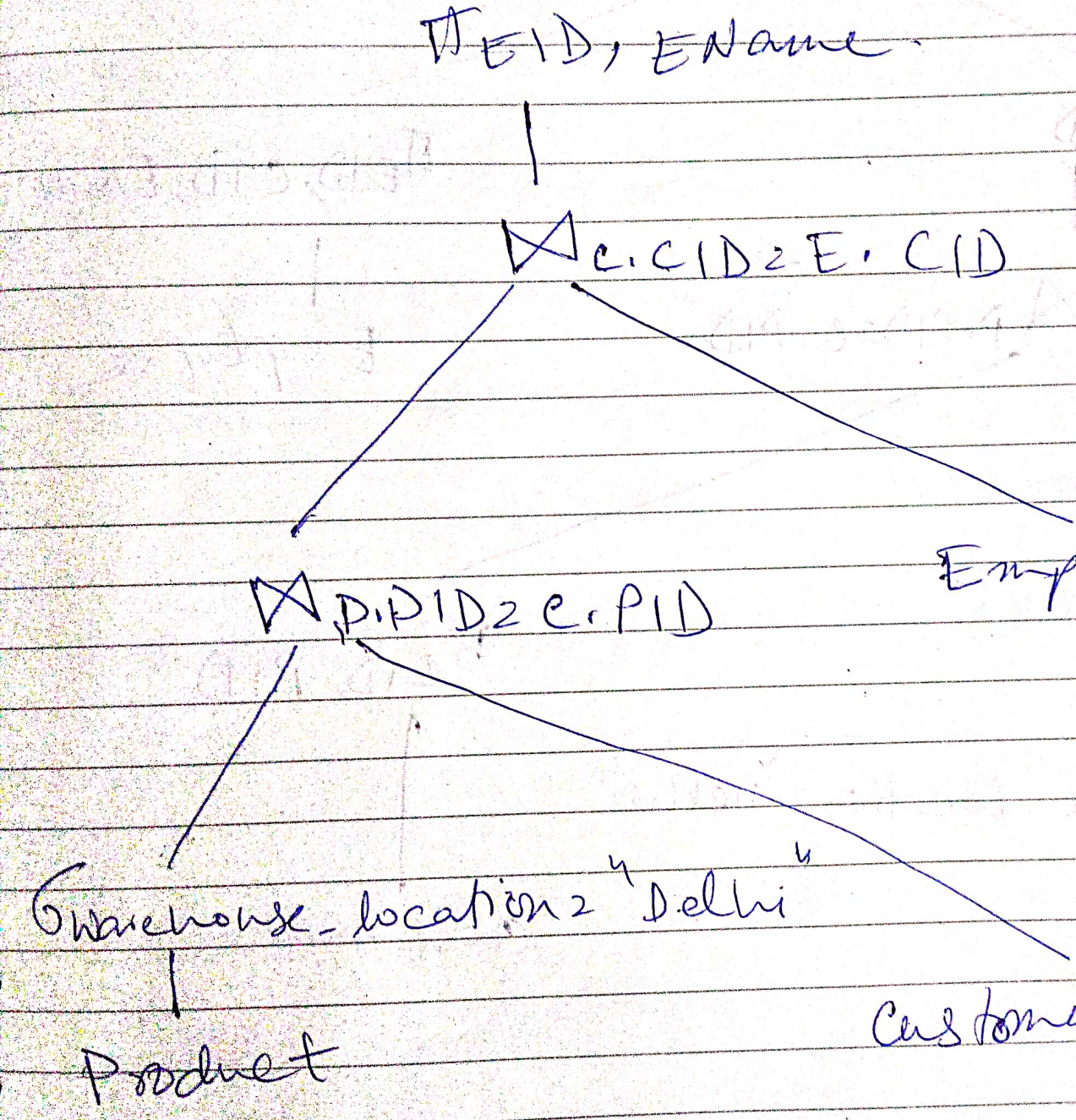
$\times$

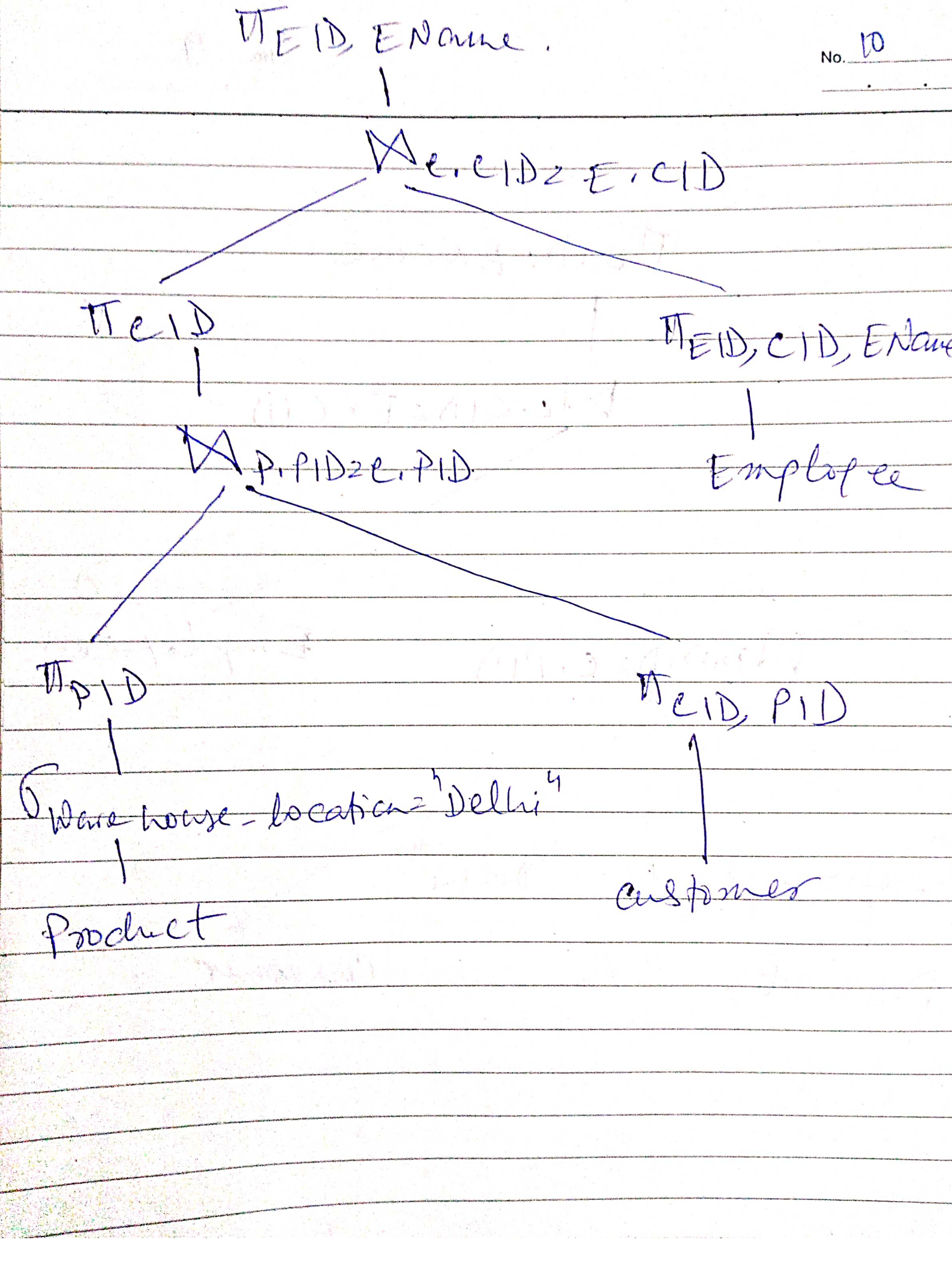
Product customer

P: Product
E: Employee
c: customer

No. 8


$\pi_{E.ID, E.Name}$
$\sigma_{c.CID = E.CID}$
$\sigma_{P.PID = c.PID}$
Employee
$\sigma_{warehouse\_location = "Delhi"}$
Product
Customer


Π EID, EName
⋈ C.CID = E.CID
Employee
⋈ P.PID = C.PID
Warehouse_location = 'Delhi'
Product
Customer


$\Pi_{EID, EName}$
$\bowtie_{e.cID=E.cID}$
Π_{cID}
$\Pi_{EID, cID, EName}$
$\bowtie_{P.PID=e.PID}$
Employee
Π_{PID}
$\Pi_{cID, PID}$
$\sigma_{Warehouse\_location='Delhi'}$
Product
customer

3ai) SELECT * FROM Television
WHERE Company-ID = 3 AND Price < 35000

3aii) Plan 1:

1. TEMP ← $\sigma_{CID=3}$.
2. $\sigma_{Price<35000}$ (TEMP).

nTuples (Temp) = r/d_{CID}

= 20000/100 = 200

nBlocks (Temp) = nTuples / BFR.

= 200/100 = 2.

Estimate using Secondary Index.

$(x_{CID} + s_{CID})$ + nblocks (Temp) + nblocks (Temp)

= (2 + 200) + 2 + 2 = 206 blocks.

Plan 2:

1. Temp ← $\sigma_{price < 35000}$.
2. $\sigma_{CID = 3}$ (Temp).

MIN price (Television) = 15000

Max price (Television) = 40000

n Tuples (Temp) = 20000 × (35000 − 15000) / (40000 − 15000)

= 16000

n blocks (Temp) = 16000/100 = 160.

Cost = $b_{Television}$ + 160 + 160

= 20000/100 + 160 + 160

= 520 blocks.

So, Plan 1 is efficient.

35i) SELECT * FROM Customer
WHERE Category_num > 5 AND.
Purchase_amt < 15000

35ii) Plan 1:

1) Temp ← $\sigma_{category\_num > 5}$.

2) $\sigma_{purchase\_amt < 15000}$ (Temp).

n Tuples (Temp).
= 15000 * (50 - 5) / (50 - 1)

Assuming max category (customer) = 50.
min category (customer) = 1.

= 13776.

n Blocks (Temp) = 13776 / 50 = 276.

Cost = (2 + 15000/50) + 276 + 276.
= 854 blocks.

Plan 2:

1) temp <- $\sigma_{\text{Purchase\_amt} < 15000}$.

2) $\sigma_{\text{Category\_num} <= 5}$ (Temp).

Min Purchase_amt (customer) = 5000.

Max Purchase_amt (customer) = 20000

n Tuples (Temp).

= 15000 * (15000 - 5000) / (20000 - 5000)

= 15000.

n Blocks (Temp) = 15000 / 50 = 300

Cost = 300 + 300 + 300

= 900 blocks.

∴ plan 1 is efficient.