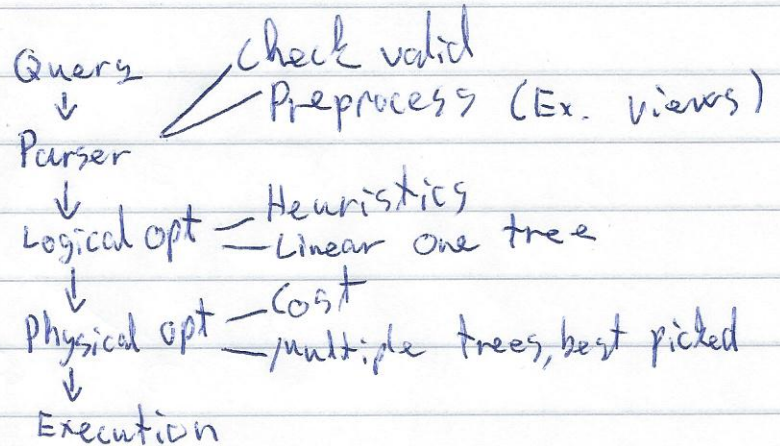


5

Query Optimization

Overview



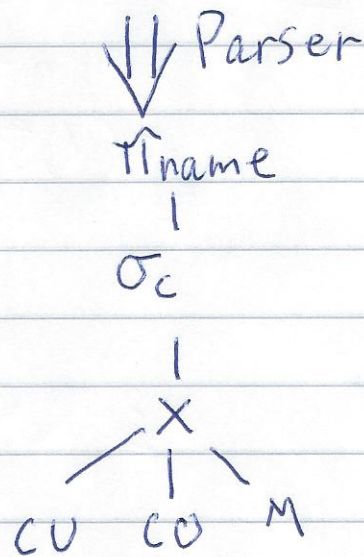
Example

"What are the names of customers living on Elm street who have checked out 'Terminator'?"

Select name

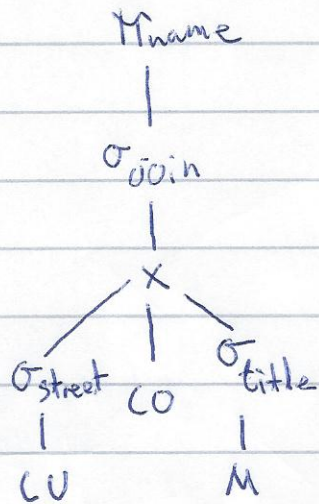
From CU, CO, M

Where CU.CID = CO.CUI and M.MID = CO.MID and CU.Street = "Elm" and M.title = "Terminator"

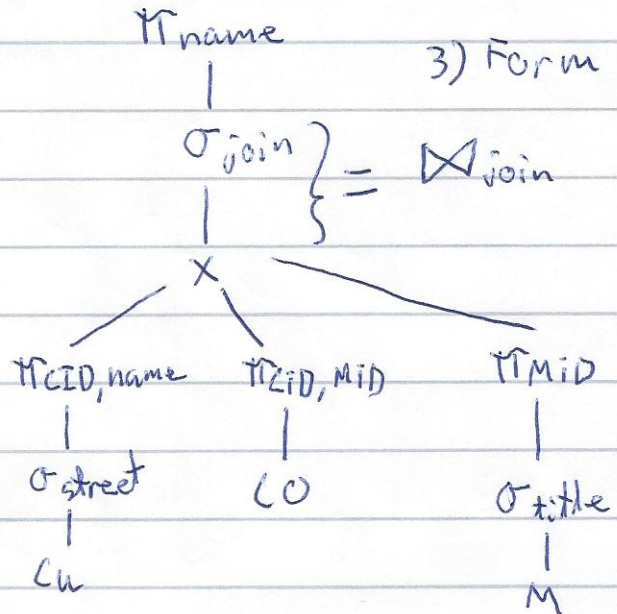


Logical optimization

1) Selections early



2) Projections early



3) Form joins

Physical optimization

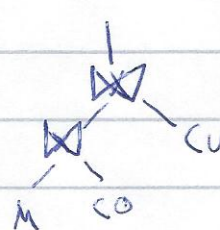
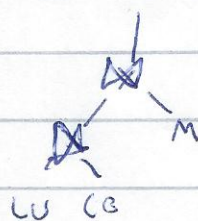
Statistics based

No. of tuples $T(R)$, Number of blocks $B(R)$

No. of different attribute values $V(R, A)$

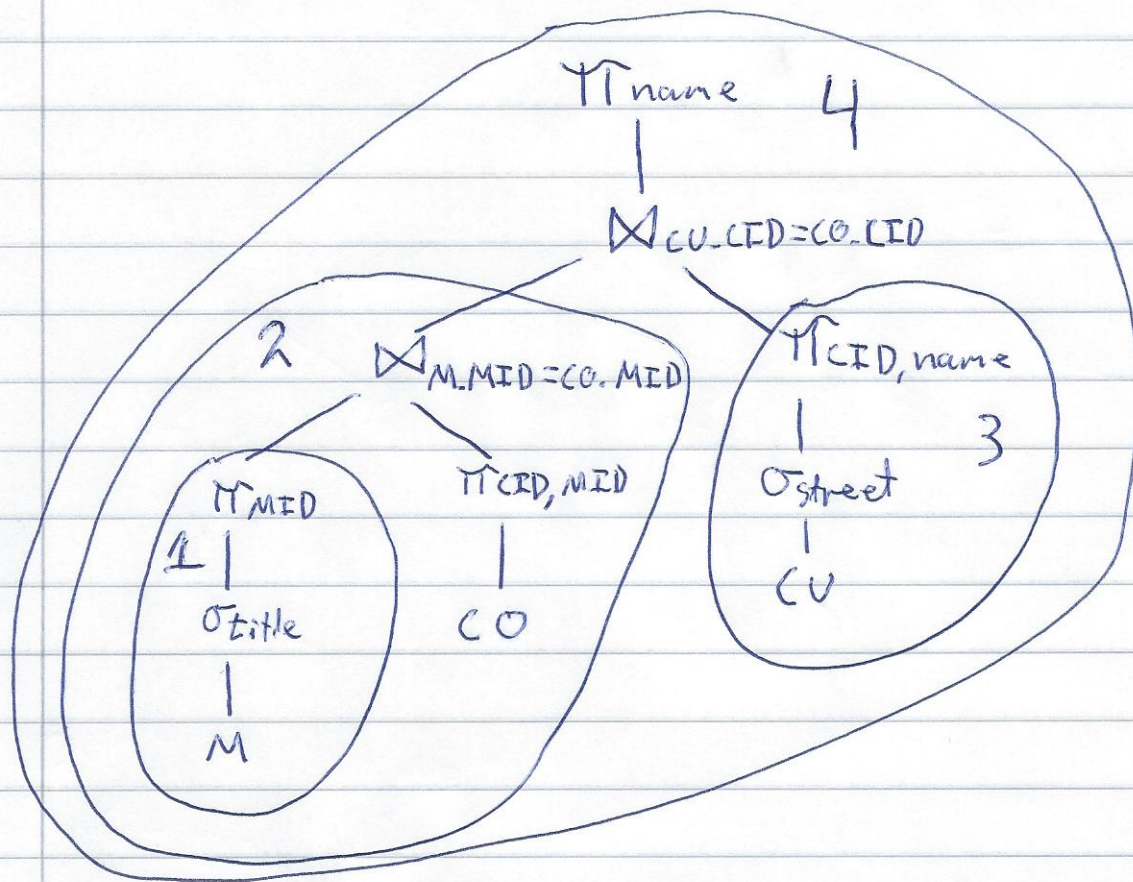
Indexes (type, levels)

Two different plans



Use most restrictive first

Calculation of cost



$$T(M) = 2000 \quad B(M) = 50 \quad T(\text{title}) = 1$$

Secondary hash index

$$\text{Cost}(1) = 1 + 1 = 2$$

$$T(CO) = 20000 \quad B(CO) = 2000 \quad T(\text{MID}) = 8$$

Secondary B-tree index with 2 levels

$$\text{Cost}(2) = 1 + 8 = 9$$

$$T(CU) = 2000 \quad B(CU) = 10 \quad T(\text{street}) = 10$$

Scan

$$\text{Cost}(3) = 10$$

$$\text{Cost}(4) = 0 \quad (\text{since in memory})$$