CSE 344

SECTION 4 – RELATIONAL ALGEBRA

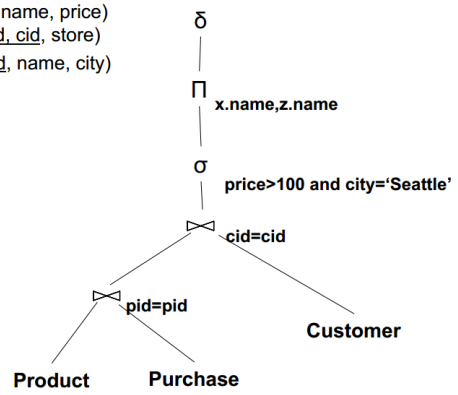
Why RA?

- Formalism for describing queries
- Basis of relational databases
- Will make you a SQL wizard!

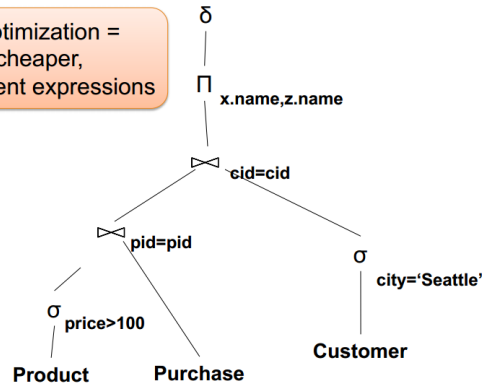
Notes on RA

Multiple possible query plans

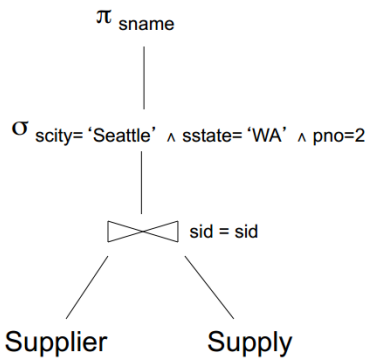
Product(pid, name, price)
Purchase(pid, cid, store)
Customer(cid, name, city)

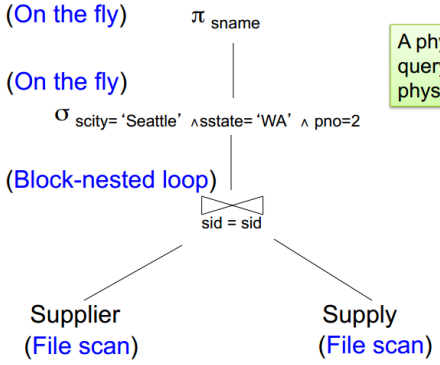


Query optimization = finding cheaper, equivalent expressions



Logical vs. Physical query plans

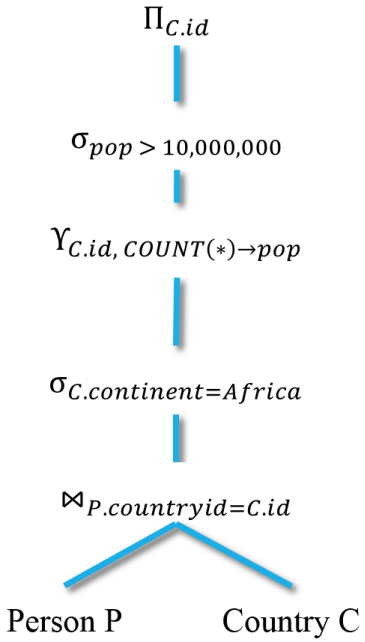




A physical query plan is a logical query plan annotated with physical implementation details

Example: RA-to-SQL

Person(id, name, countryid)
Country(id, name, continent)




```
SELECT C.id
FROM Person P, Country C
WHERE P.countryid = C.id
AND C.continent='Africa'
GROUP BY C.id
HAVING COUNT(*) > 10000000
```

Can we make a more efficient plan?

Equivalently in equation form: $\Pi_{C.id}(\sigma_{pop>10{,}000{,}00}(\Upsilon_{C.id,COUNT(*)\rightarrow pop}(\sigma_{C.continent=Africa}(Person\ P \bowtie_{P.countryid=C.id} Country\ C))))$

Demo in Azure!

RA Reference Sheet

Name	Symbol
Selection	σ
Projection	π
Join	$\bowtie$
Group By	γ
Set Difference	$-$
Duplicate Elimination	δ