ADVANCED TOPICS IN DATABASES

Relational algebra
Exercise

Master in Informatics Engineering
Data Engineering

Informatics Engineering Department

ISEP INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO

Renaming

- **Renaming -**

Helper operator ρ does not derive new result, just renames relations and fields

The renaming operation is represented by the expressions

ρS(R) or ρ(B1, B2, ..., Bn)(R) or ρS(B1, B2, ..., Bn)(R)

where ρ is the renaming operator, S is the new relation name and B1, B2, ..., Bn are the new attribute names

$$DEP4\_SAL2000 \leftarrow \sigma_{NumDep = 4 \text{ AND } Salário > 2000}(EMPREGADO)$$
$$RESULT \leftarrow \pi_{NumBI,\ NomeP,\ NomeF}(DEP4\_SAL2000)$$

$$\rho_{DEP4\_SAL2000}(\sigma_{NumDep = 4 \text{ AND } Salário > 2000}(EMPREGADO))$$
$$\rho_{RESULT(BI,\ Nome,\ Apelido)}(\pi_{NumBI,\ NomeP,\ NomeF}(DEP4\_SAL2000$$

RESULT	BI	Nome	Apelido
	798764544	João	Santos
	342342324	Ana	Feio

Grouping and Aggregation

Find name of oldest sailor - only basic relational álgebra operators

sailors

sid	sname	rating	age
22	dustin	7	45.0
31	lubber	8	55.5
58	rusty	10	35.0

1. Renaming sailors

$\rho_{\text{sailor1}}(\text{sailors})$

sailor1

sid	sname	rating	age
22	dustin	7	45.0
31	lubber	8	55.5
58	rusty	10	35.0

2 .Cartesian Product

sailor x sailor1

(sid)	sname	rating	age	(sid)	sname	rating	age
22	dustin	7	45	22	dustin	7	45.0
22	dustin	7	45	31	lubber	8	55.5
22	dustin	7	45	58	rusty	10	35.0
31	lubber	8	55.5	22	dustin	7	45.0
31	lubber	8	55.5	31	lubber	8	55.5
31	lubber	8	55.5	58	rusty	10	35.0
58	rusty	10	35.0	22	dustin	7	45.0
58	rusty	10	35.0	31	lubber	8	55.5
58	rusty	10	35.0	58	rusty	10	35.0

Grouping and Aggregation

3. Find the youngest sailor Renaming sailors

$$\rho \text{ youngest } (\pi\ sailors.sname, sailors.age\ (\sigma_{\ sailors.\ age\ <\ sailor1.\ age}\ (sailors\ x\ sailor1)$$

4. Difference sailors with youngest.

$\pi\ sailors.sname, sailors.age$ (Sailors) – youngest

(sid)	sname	rating	age	(sid)	sname	rating	age
22	dustin	7	45	22	dustin	7	45.0
22	dustin	7	45	31	lubber	8	55.5
22	dustin	7	45	58	rusty	10	35.0
31	lubber	8	55.5	22	dustin	7	45.0
							55.5
							35.0
							45.0
							55.5
							35.0

(sid)	s
31	l

$\rho_{\ \text{sailor1}}$ (sailors)

ρ youngest $(\pi\ sailors.sname, sailors.age\ (\sigma_{\ sailors.\ age\ <\ sailor1.\ age}\ (sailors\ x\ sailor1)$

Sailors – Youngest

Grouping and Aggregation

- Very useful to apply a function to a collection of values to generate a single result
- Most common aggregate functions:

  sum sums the values in the collection

  avg computes average of values in the collection

  count counts number of elements in the collection

  min returns minimum value in the collection

  max returns maximum value in the collection

- Aggregate functions work on multisets, not sets
- A value can appear in the input multiple times

$$\gamma_{\text{sname,max (age)}}(\text{Sailors})$$

```
Select age, sum(rating)
From Sailors
Group by age
Having count(*) >50
```

$$\gamma_{\text{age, sum (rating),count(*)>50}}(\text{Sailors})$$

Grouping and Aggregation

"Find the total amount owed to the credit company."

$$\mathcal{G}_{\mathbf{sum}(balance)}(credit\_acct)$$

4275

cred_id	limit	balance
C-273	2500	150
C-291	750	600
C-304	15000	3500
C-313	300	25

credit_acct

"Find the maximum available credit of any account."

$$\mathcal{G}_{\mathbf{max}(available\_credit)}(\Pi_{(limit - balance)\ \mathbf{as}\ available\_credit}(credit\_acct))$$

11500

Exercise 1

A. Given the relations that are presented below, which expressions using only R and S allow us to construct the relations T and U?

R

A	B	C
a	b	c
d	b	c
b	b	f
c	a	d

S

B	C	D
b	c	d
b	c	e
a	d	b

T

A	B	C	D
a	b	c	d
a	b	c	e
d	b	c	d
d	b	c	e
c	a	d	b

U

A	D
a	d
a	e
d	d
d	e

Exercise 2

B. (Exercise 5.2.9, J. Ullman and J. Widom)

LAPTOP

model	speed	ram	hd	Scree	price
104	2.66	1024	240	20.1	3245
105	1.50	512	120	15.3	1211
107	2.00	2048	100	15.4	890
112	2.00	1024	60	17.0	2123
113	1.60	2048	60	17.0	2300

PRINTER

model	color	type	price
110	true	Ink-jet	99
111	false	Laser	234
114	true	Laser	890
115	true	Ink-jet	120

PC

Model	Speed	Ram	Hd	Price
101	2.66	1024	250	1234
102	1.50	8192	250	1211
103	2.80	2048	100	1342
106	1.86	1024	100	890
108	3.06	8192	80	956
109	2.03	1024	120	1154
116	2.06	2048	512	1211
117	3.50	1024	120	1423

PRODUCT

maker	model	type
A	101	PC
A	102	PC
A	103	PC
A	104	LAPTOP
B	105	LAPTOP
B	106	PC
C	107	LAPTOP
D	108	PC
E	109	PC
E	110	PRINTER
F	111	PRINTER
F	112	LAPTOP
F	113	LAPTOP
G	114	PRINTER
G	115	PRINTER
D	116	PC
D	117	PC

a) What PC models have a speed of at least 2.2 MHz?

$\pi_{model}(\sigma_{speed \geq 2.2}(PC))$

b) Which manufacturers make laptops with a hard disk (hd) of at least 100 GB?

$\pi_{maker}(\sigma_{hd>100}(PRODUCT \bowtie LAPTOP))$

b) Find the model number and price of all PC's made by manufacturer B.

$\pi_{model,price}(PC \bowtie (\sigma_{maker='B'}(PRODUCT)))$

d) Find those manufacturers that sell Laptops, but not PC's.

$\pi_{maker}(PRODUCT \bowtie (\pi_{model}(LAPTOP) - \pi_{model}(PC)))$

e) Find those hard disk(hd) sizes that occur in two or more PCs

$PC1 = \pi_{md,hd1}(\rho_{md \leftarrow model, sp \leftarrow speed, hd1 \leftarrow hd} PC)$

$PC2 = \pi_{md1,hd2}(\rho_{md1 \leftarrow model, sp1 \leftarrow speed, hd2 \leftarrow hd} PC)$

$PC1 \bowtie_{(PC.md = PC.md1 \wedge PC.hd1 = PC.hd2)} PC2$

Exercise 3

Convert to relational algebra

First

```
SELECT agent_code, COUNT(agent_code) mycount
 FROM orders
 GROUP BY agent_code);
```

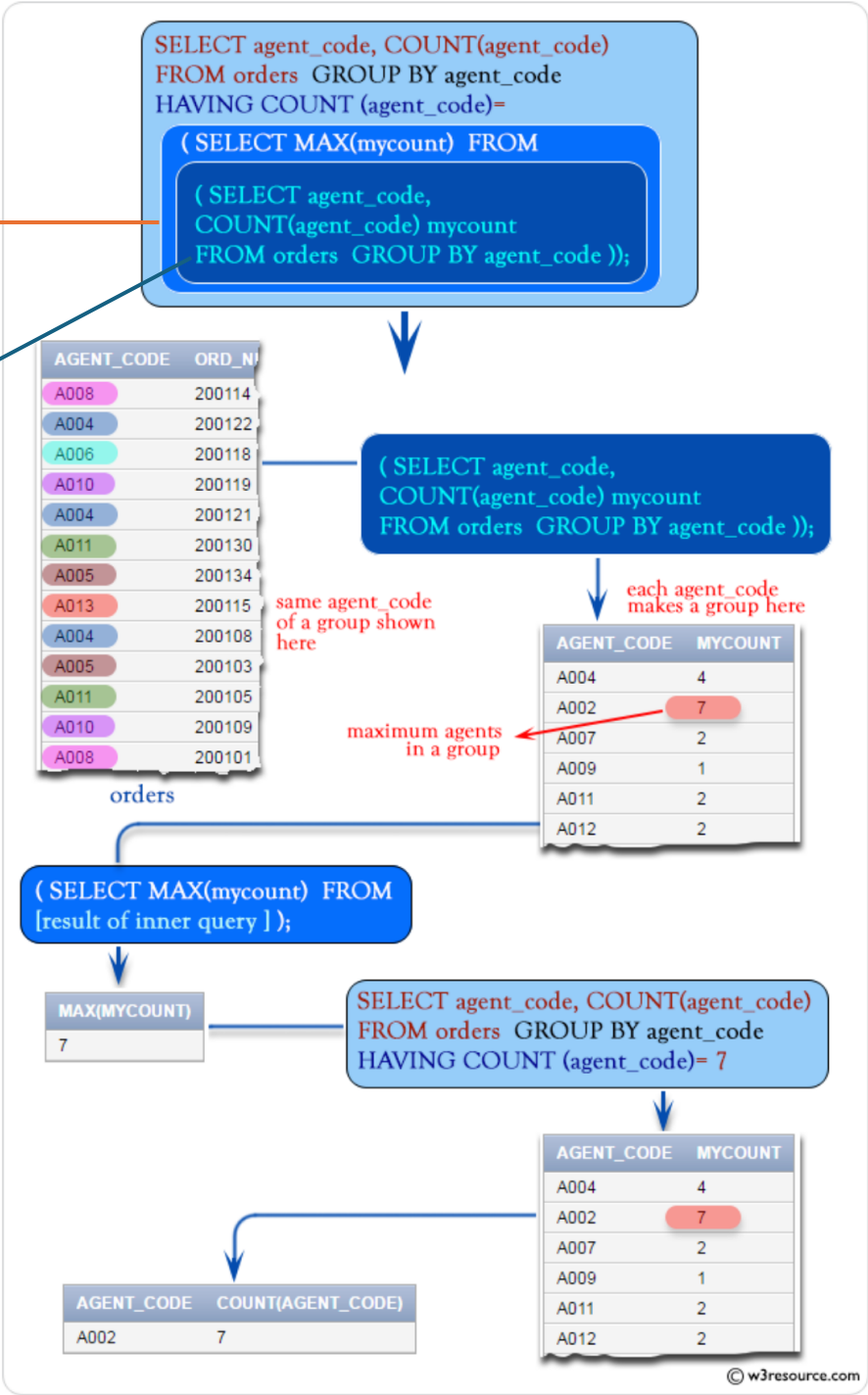
$(\gamma_{\text{agent\_code, COUNT(agent\_code) -> mycount}}(\text{orders}))$

Next :

```
SELECT MAX( mycount) FROM
        (SELECT agent_code, COUNT(agent_code) mycount
          FROM orders
          GROUP BY agent_code));
```

$\gamma_{\text{MAX(mycount) -> mmc}}(\gamma_{\text{agent\_code, COUNT(agent\_code) -> mycount}}(\text{orders}))$

https://dbis-uibk.github.io/relax/landing



Exercise 4

Write an expression that is equivalent to the following SQL query:

```
SELECT  team_id AS tid
FROM players
WHERE players.team_id  NOT IN
                              (SELECT  team_id  FROM teams )
AND position = “shooting_guard”
```

$$P_{teamid} \leftarrow tid(\pi_{teamid}(\sigma_{position='shooting\ guard'}(players)) \quad -$$

$$\pi_{players.teamid}(players \bowtie_{players.teamid=teams.teamid} teams))$$