



Università degli Studi di Padova

## **Relational Algebra**

### Basi di Dati

Bachelor's Degree in Computer Engineering Academic Year 2024/2025



#### **Ornella Irrera**

Intelligent Interactive Information Access (IIIA) Hub Department of Information Engineering University of Padua





Given two relations  $R_1(X)$  and  $R_2(Y)$  and a proposition  $\Theta$ , the theta-join is the relation  $R_1 \Join_{\Theta} R_2 = \{t \mid \text{exists } x \in R_1, y \in R_2 \text{ such that} t[X] = x \land t[Y] = y \land \Theta\}$ 

Given two relations  $R_1(X_1, X_2, \ldots, X_n)$  and  $R_2(Y_1, Y_2, \ldots, Y_m)$ , the theta-join  $Q = R_1 \Join_{\Theta} R_2 = (X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m)$  has degree q = n + m

• Let  $|R_1| = n_{R_1}$  and  $|R_2| = n_{R_2}$  be the cardinalities of the two relations, the cardinality of the thetajoin, also called **selectivity**, is

$$0 \le |Q| = n_Q \le n_{R_1} * n_{R_2}$$

The proposition  $\Theta$  can be defined as follows

- $X_i \ \theta \ Y_i$  with  $X_i$  and  $Y_i$  attributes of  $R_1$  and  $R_2$  on the same domain and  $\theta \in \{<, >, =, \neq, \leq, \geq\}$  comparison operator
- if  $\phi$  and  $\psi$  are propositions, then also  $\phi \wedge \psi$  is a proposition





Given two relations  $R_1(X)$  and  $R_2(Y)$  and a proposition O  $R_1 \bowtie_{\Theta} R$  $x \in R_1, y \in R_2$  such that Equivalence Equivalence  $R_1 \times R_2$   $R_1 \times R_2 = \sigma \Theta(R_1 \times R_2)$   $R_1 \times \Theta R_2 = \sigma \Theta(R_1 \times R_2)$   $R_1 \times \Theta R_2$   $R_1 \times \Theta R_2$ Given two relations Rtheta-join has degree q = n + m• Let  $|R_1| = n_{R_1}$  and  $|R_2|$ lations, the cardinality of the thetajoin, also called **selectivity**,

The proposition  $\Theta$  can be defined as follows

- $X_i \ \theta \ Y_i$  with  $X_i$  and  $Y_i$  attributes of  $R_1$  and  $R_2$  on the same domain and  $\theta \in \{<, >, =, \neq, \leq, \geq\}$  comparison operator
- if  $\phi$  and  $\psi$  are propositions, then also  $\phi \wedge \psi$  is a proposition





#### Graduated

GBadge	GSurname	GAge
7274	Rossi	42
7432	Neri	54
9824	Verdi	45

#### Manager

MBadge	MSurname	MAge
9297	Neri	33
7432	Neri	54
9824	Verdi	45

a a a a a a a a a a a a a a a a a a a		GBadge	GSurname	GAge	MBadge	MSurname	MAge
---------------------------------------	--	--------	----------	------	--------	----------	------



### Product (cartesian): example



#### Graduated

GBadge	GSurname	GAge
7274	Rossi	42
7432	Neri	54
9824	Verdi	45

#### Graduated × Manager

#### Manager

MBadge	MSurname	MAge
9297	Neri	33
7432	Neri	54
9824	Verdi	45

GBadge	GSurname	GAge	MBadge	MSurname	MAge
7274	Rossi	42	9297	Neri	33
7274	Rossi	42	7432	Neri	54
7274	Rossi	42	9824	Verdi	45
7432	Neri	54	9297	Neri	33
7432	Neri	54	7432	Neri	54
7432	Neri	54	9824	Verdi	45
9824	Verdi	45	9297	Neri	33
9824	Verdi	45	7432	Neri	54
9824	Verdi	45	9824	Verdi	45

Basi di Dati, A.Y. 2024/2025 BD in "Computer Engineering"





#### Graduated

GBadge	GSurname	GAge
7274	Rossi	42
7432	Neri	54
9824	Verdi	45

#### Manager

MBadge	MSurname	MAge
9297	Neri	33
7432	Neri	54
9824	Verdi	45

#### Graduated M GAge > MAge Manager

GBadge	GSurname	GAge	MBadge	MSurname	MAge
7274	Rossi	42	9297	Neri	33
7432	Neri	54	9297	Neri	33
7432	Neri	54	9824	Verdi	45
9824	Verdi	45	9297	Neri	33





#### Project

#### Employee

Location	Pnumber	Hours	Pname
Padova	1	32.5	Product X
Padova	2	7.5	Product Y
Roma	2	20.0	Product Y

SSNEmp	EnumberPrj	ELocation	Ename
123456	1	Padova	Smith John
236711	2	Padova	Doe John
453453	2	Roma	English Joyce





#### Project

#### Employee

Location	Pnumber	Hours	Pname
Padova	1	32.5	Product X
Padova	2	7.5	Product Y
Roma	2	20.0	Product Y

SSNEmp	EnumberPrj	ELocation	Ename
123456	1	Padova	Smith John
236711	2	Padova	Doe John
453453	2	Roma	English Joyce

#### **Project x Employee**

Location	Pnumber	Hours	Pname	SSNEmp	ELocation	EnumberPro	Ename
Padova	1	32.5	Product X	123456	Padova	1	Smith John
Padova	1	32.5	Product X	236711	Padova	2	Doe John
Padova	1	32.5	Product X	453453	Roma	2	English Joyce
Padova	2	7.5	Product Y	123456	Padova	1	Smith John
Padova	2	7.5	Product Y	236711	Padova	2	Doe John
Padova	2	7.5	Product Y	453453	Roma	2	English Joyce
Roma	2	20.0	Product Y	123456	Padova	1	Smith John
Roma	2	20.0	Product Y	236711	Padova	2	Doe John
Roma	2	20.0	Product Y	453453	Roma	2	English Joyce





#### Project Pnumber Location Hours Pname Product X Padova 32.5 1 Product Y Padova 2 7.5 2 20.0 Product Y Roma

SSNEmp	EnumberPrj	ELocation	Ename
123456	1	Padova	Smith John
236711	2	Padova	Doe John
453453	2	Roma	English Joyce

Employee

#### **Project** ⋈ Location = ELocation AND PNumber = EnumberProj **Employee**

Location	Pnumber	Hours	Pname	SSNEmp	ELocation	EnumberPro j	Ename
Padova	1	32.5	Product X	123456	Padova	1	Smith John
Padova	2	7.5	Product Y	236711	Padova	2	Doe John
Roma	2	20.0	Product Y	453453	Roma	2	English Joyce





#### Project

#### Employee

Location	Pnumber	Hours	Pname
Padova	1	32.5	Product X
Padova	2	7.5	Product Y
Roma	2	20.0	Product Y

SSNEmp	EnumberPrj	ELocation	Ename
123456	1	Padova	Smith John
236711	2	Padova	Doe John
453453	2	Roma	English Joyce





#### Project

#### Employee

Location	Pnumber	Hours	Pname
Padova	1	32.5	Product X
Padova	2	7.5	Product Y
Roma	2	20.0	Product Y

SSNEmp	EnumberPrj	ELocation	Ename
123456	1	Padova	Smith John
236711	2	Padova	Doe John
453453	2	Roma	English Joyce

#### **Project x Employee**

Location	Pnumber	Hours	Pname	SSNEmp	ELocation	EnumberPro	Ename
Padova	1	32.5	Product X	123456	Padova	1	Smith John
Padova	1	32.5	Product X	236711	Padova	2	Doe John
Padova	1	32.5	Product X	453453	Roma	2	English Joyce
Padova	2	7.5	Product Y	123456	Padova	1	Smith John
Padova	2	7.5	Product Y	236711	Padova	2	Doe John
Padova	2	7.5	Product Y	453453	Roma	2	English Joyce
Roma	2	20.0	Product Y	123456	Padova	1	Smith John
Roma	2	20.0	Product Y	236711	Padova	2	Doe John
Roma	2	20.0	Product Y	453453	Roma	2	English Joyce





#### Project

#### Employee

Location	Pnumber	Hours	Pname	SSNEmp	EnumberPrj	ELocation	Ename
Padova	1	32.5	Product X	123456	1	Padova	Smith John
Padova	2	7.5	Product Y	236711	2	Padova	Doe John
Roma	2	20.0	Product Y	453453	2	Roma	English Joyce

#### **Project** ⋈ Location = ELocation OR PNumber = EnumberProj Employee

Location	Pnumber	Hours	Pname	SSNEmp	ELocation	EnumberPro	Ename
Padova	1	32.5	Product X	123456	Padova	1	Smith John
Padova	1	32.5	Product X	236711	Padova	2	Doe John
Padova	2	7.5	Product Y	123456	Padova	1	Smith John
Padova	2	7.5	Product Y	236711	Padova	2	Doe John
Padova	2	7.5	Product Y	453453	Roma	2	English Joyce
Roma	2	20.0	Product Y	236711	Padova	2	Doe John
Roma	2	20.0	Product Y	453453	Roma	2	English Joyce





#### Project

#### Employee

Location	Pnumber	Hours	Pname	SSNEmp	EnumberPrj	ELocation	Ename
Padova	1	32.5	Product X	123456	1	Padova	Smith John
Padova	2	7.5	Product Y	236711	2	Padova	Doe John
Roma	2	20.0	Product Y	453453	2	Roma	English Joyce

#### **Project** ⋈ Location = ELocation OR PNumber = EnumberProj Employee

Location	Pnumber	Hours	Pname	SSNEmp	ELocation	EnumberPro	Ename
Padova	1	32.5	Product X	123456	Padova	1	Smith John
Padova	1	32.5	Product X	236711	Padova	2	Doo John
Padova	2	7.5	Product Y	123456	Padova	1	Smith John
Padova	2	7.5	Product Y	236711	OR produ Padova	ices spuriou	s tuples!! Doe John
Padova	2	7.5	Product Y	453453	Roma	2	English Joyce
Roma	2	20.0	Product Y	236711	Padova	2	Doe John
Roma	2	20.0	Product Y	453453	Roma	2	English Joyce





### Project

#### Employee

Location	Pnumber	Hours	Pname	SSNEmp	EnumberPrj	ELocation	Ename
Padova	1	32.5	Product X	123456	1	Padova	Smith John
Padova	2	7.5	Product Y	236711	2	Padova	Doe John
Roma	2	20.0	Product Y	453453	2	Roma	English Joyce

#### **Project** ⋈ Location = ELocation OR PNumber = EnumberProj Employee

Location	Pnumber	Hours	Pname	SSNEmp	ELocation	EnumberPro	Ename
Padova	1	32.5	Product X	123456	Padova	1	Smith John
Padova	1	32.5	Product X	236711	Padova	2	Doo John
Padova	2	7.5	Popular	123 <mark>456</mark>	Padova	1	Smith John
Padova	2	7.5	Pro LCLY	236711	OR produ Padova	ices spuriou	s tuples!! Doe John
Padova	2	7.5	Product Y	453 <mark>453</mark>	Roma	2	English Joyce
Roma	2	20.0	Product Y	236711	Padova	2	Doe John
Roma	2	20.0	Product Y	453453	Roma	2	English Joyce





## The equi-join is a theta-join where we can use only the = comparison operator





# The equi-join is a theta-join where we can use only the = comparison operator

#### **Employee**

Surname	Dep
Rossi	A
Neri	В
Bianchi	B

#### Department

Code	Manager
A	Mori
B	Bruni

#### **Employee** $\bowtie_{Dep = Code}$ **Department**

Surname Dep Code Manager





# The equi-join is a theta-join where we can use only the = comparison operator



#### **Employee** $\bowtie_{Dep = Code}$ **Department**

Surname	Dep	Code	Manager
Rossi	A	A	Mori
Neri	B	B	Bruni
Bianchi	B	B	Bruni

Basi di Dati, A.Y. 2024/2025 BD in "Computer Engineering"





The natural join is an equi-join where the attributes of the two relations in the join condition have the same names and the duplicated attributes are removed from the schema of the output relation

• Let  $R_1(XW)$  and  $R_2(YZ)$  the two relations with W and Z sets of attributes with the same name, then the natural join can be expressed as

 $R_1 \bowtie R_2 = \pi_{X, W_1, W_2, \dots, W_m, Y} (R_1 \bowtie_{W_1 = Z_1 \land W_2 = Z_2 \land \dots \land W_m = Z_m} R_2)$ 



Surname	Dep
Rossi	A
Neri	B
Bianchi	B

#### Department

Dep	Manager
A	Mori
B	Bruni

#### Employee 🛛 Department

Surname Dep Manager





#### **Employee** $\bowtie$ **Department**

Surname	Dep	Manager
Rossi		Mori
Neri	B	Bruni
Bianchi	B	Bruni



#### Employee (R<sub>1</sub>)

Surname	Dep
Rossi	A
Neri	B
Bianchi	B
Neri	С

#### Department (R<sub>2</sub>)



#### **Employee** $\bowtie$ **Department**

Surname	Dep	Manager
Rossi		Mori
Neri	В	Bruni
Neri	B	Verdi
Bianchi	B	Bruni
Bianchi	В	Verdi

Basi di Dati, A.Y. 2024/2025 BD in "Computer Engineering"



### Cardinality of the Natural Join: Example (1/3)



Basi di Dati, A.Y. 2024/2025 BD in "Computer Engineering"

© Ornella Irrera



### Employee (R<sub>1</sub>)

Surname	Dep
Rossi	A
Neri	B
Bianchi	В
Neri	С

#### **Department (R<sub>2</sub>)**



#### **Employee** $\bowtie$ **Department**

Surname	Dep	Manager
Rossi		Mori
Neri	В	Bruni
Bianchi	B	Bruni



### Cardinality of the Natural Join: Example (2/3)





### Cardinality of the Natural Join: Example (3/3)



#### **Employee** $\bowtie$ **Department**

Surname	Rep	Manager
Rossi		Mori
Neri	В	Bruni
Bianchi	B	Bruni
Neri		Verdi



### Cardinality of the Natural Join: Example (3/3)



## Cardinality of the Natural Join (Equi-Join)

As in the general case, the cardinality of the natural join (equi-join) is

$$0 \le |R_1 \bowtie R_2| \le |R_1| * |R_2|$$

If the natural join (equi-join) involves attributes K which are a key (maybe primary, PK) of  $R_2$  then

 $0 \le |R_1 \bowtie R_2| \le |R_1|$ 

If the natural join (equi-join) involves attributes K which are a key (maybe primary, PK) of  $R_2$  and the same attributes are a foreign key FK of  $R_1$  with a referential integrity constraint on  $R_2$  then

$$R_1 \bowtie R_2 | = |R_1|$$



Badge	Surname	Age	Branch
7309	Rossi	34	Roma
5998	Bianchi	42	Milano
9553	Neri	NULL	Milano



Badge	Surname	Age	Branch
7309	Rossi	34	Roma
5998	Bianchi	42	Milano
9553	Neri	NULL	Milano

### $\sigma_{Age>40}(Employee)$

Badge	Surname	Age	Branch
5998	Bianchi	42	Milano



Badge	Surname	Age	Branch
7309	Rossi	34	Roma
5998	Bianchi	42	Milano
9553	Neri	NULL	Milano

### $\sigma_{\text{Age}>40}(\text{Employee}) \cup \sigma_{\text{Age}\leq40}(\text{Employee}) \neq \text{Employee}$

Selections are evaluated separately



Badge	Surname	Age	Branch
7309	Rossi	34	Roma
5998	Bianchi	42	Milano
9553	Neri	NULL	Milano

 $\sigma_{\text{Age}>40}$   $\swarrow_{\text{Age}\leq40}$  (Employee)  $\neq$  Implegato

Atomic conditions are evaluated separately



# To manage NULL values we introduce specific conditions: IS NULL Employee IS NOT NULL Badge Surname Age Branch

Badge	Surname	Age	Branch
7309	Rossi	34	Roma
5998	Bianchi	42	Milano
9553	Neri	NULL	Milano

#### $\sigma_{Age}$ is NULL (Employee)

Badge	Surname	Age	Branch
9553	Neri	NULL	Milano

### $\sigma_{\rm Age}\,{\rm is \, not\, null}({\rm Employee})$

Badge	Surname	Age	Branch
7309	Rossi	34	Roma
5998	Bianchi	42	Milano



### Managing the NULL Values: Example

#### Employee

Badge	Surname	Age	Branch
7309	Rossi	34	Roma
5998	Bianchi	42	Milano
9553	Neri	NULL	Milano

 $\sigma_{\text{Age}>40 \vee \text{Age}\leq40 \vee \text{Age}}$  IS NULL (Employee) = Employee

$$\sigma_{Age>40}(Employee)$$

$$\cup$$

$$\sigma_{Age\leq40}(Employee)$$

$$\cup$$

$$\sigma_{Age IS NULL}(Employee)$$

Basi di Dati, A.Y. 2024/2025 BD in "Computer Engineering"



### **Join: Dangling Tuples**



#### **Employee (R<sub>1</sub>)**

Surname	Dep
Rossi	A
Neri	B
Bianchi	В
Neri	C

#### **Department (R<sub>2</sub>)**



#### **Employee** $\bowtie$ **Department**

Surname	Dep	Manager
Rossi		Mori
Neri	B	Bruni
Bianchi	B	Bruni

Some tuples are dangling and do not contribute to the

#### results

Basi di Dati, A.Y. 2024/2025 BD in "Computer Engineering"





### Given two relations $R_1(XW)$ and $R_2(YZ)$ with W and Zsets of attributes on the same domain, the full outer join is

 $R_1 \bowtie_W \theta_Z R_2 = R_1 \bowtie_W \theta_Z R_2$ 

 $\bigcup_{\substack{(R_1 - \pi_{X,W}(R_1 \bowtie_{W \theta Z} R_2)) \times \{Y = \mathsf{NULL}, Z = \mathsf{NULL}\}} \\ \bigcup_{\substack{(R_2 - \pi_{Y,Z}(R_1 \bowtie_{W \theta Z} R_2)) \times \{X = \mathsf{NULL}, W = \mathsf{NULL}\}} \\ \end{array}$ 

The full outer join extends, with NULL values, the tuples which are excluded from the theta-join, also called inner join, keeping tuples from both operands









 $R_1 \bowtie_W \theta_Z R_2 = R_1 \bowtie_W \theta_Z R_2$ 

 $(R_2 - \pi_{Y,Z}(R_1 \bowtie_W \theta_Z R_2)) \times \{X = \mathsf{NULL}, W = \mathsf{NULL}\}$ 



#### Employee (R<sub>1</sub>)

Surname	Dep
Rossi	A
Neri	В
Bianchi	В
Neri	C

#### Department (R<sub>2</sub>)



#### Employee Depertment

Surname	Dep	Code	Manager
Rossi	A	A	Mori
Neri	B	B	Bruni
Bianchi	B	B	Bruni
Neri	С	NULL	NULL
NULL	NULL		Verdi



### **Full Outer Join: Example**





### **Full Outer Join: Example**





### **Full Outer Join: Example**



Basi di Dati, A.Y. 2024/2025 BD in "Computer Engineering"



### Left Outer Join: Example

#### Employee (R<sub>1</sub>)

Surname	Dep
Rossi	A
Neri	B
Bianchi	В
Neri	C

#### Department (R<sub>2</sub>)



#### Employee Depecode Department

Surname	Dep	Code	Manager
Rossi	A	A	Mori
Neri	B	B	Bruni
Bianchi	B	B	Bruni
Neri	С	0 NULL	NULL



### Left Outer Join: Example





### Left Outer Join: Example





#### Employee (R<sub>1</sub>)

Surname	Dep
Rossi	A
Neri	В
Bianchi	В
Neri	C

#### **Department (R<sub>2</sub>)**



#### Employee McDep=Code Department

Surname	Dep	Code	Manager
Rossi	A	A	Mori
Neri	B	B	Bruni
Bianchi	B	B	Bruni
NULL	NULL		Verdi



### **Right Outer Join: Example**





### **Right Outer Join: Example**







#### Employee (R<sub>1</sub>)

Surname	Dep
Rossi	A
Neri	В
Bianchi	В
Neri	NULL

#### Department (R<sub>2</sub>)



#### Employee Depertment

Surname	Dep	Code	Manager
Rossi	A	A	Mori
Neri	B	B	Bruni
Bianchi	B	B	Bruni
Neri	NULL	NULL	NULL
NULL	NULL	NULL	Verdi

