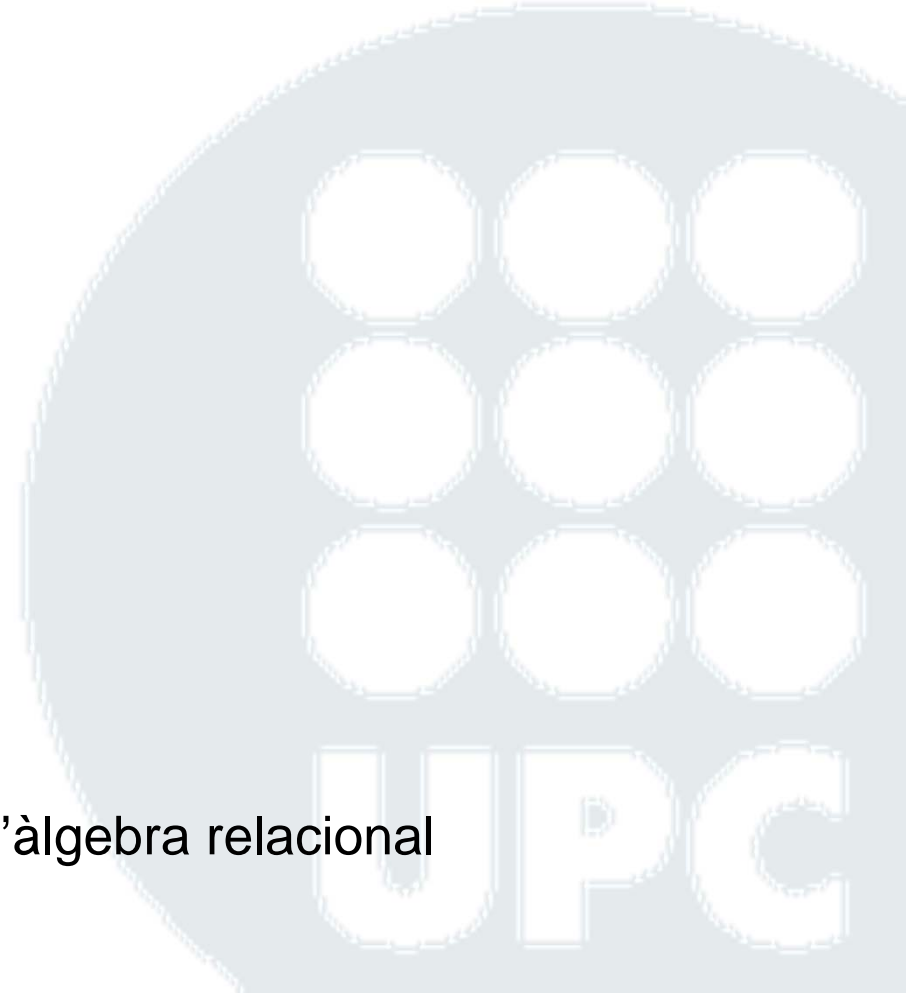


3.1 Llenguatges: Àlgebra relacional

- Introducció
- Operacions de l'àlgebra relacional
 - Unió
 - Reanomenament
 - Intersecció
 - Diferència
 - Producte cartesià
 - Selecció
 - Projectió
 - Combinació (join)
- Seqüència d'operacions de l'àlgebra relacional



Introducció

- Els **llenguatges relacionals de manipulació** (DML) es poden classificar en:

- llenguatges basats en l'àlgebra relacional
- llenguatges basats en el càlcul relacional (p.e.: SQL)

encara que molts agafen elements de totes dues línies (SQL també incorpora elements de l'àlgebra).

- **Càlcul relacional**: Té el seu fonament en el càlcul de predicats
- **Àlgebra relacional**: Té el seu fonament en la teoria de conjunts (recordar que les relacions són conjunts)

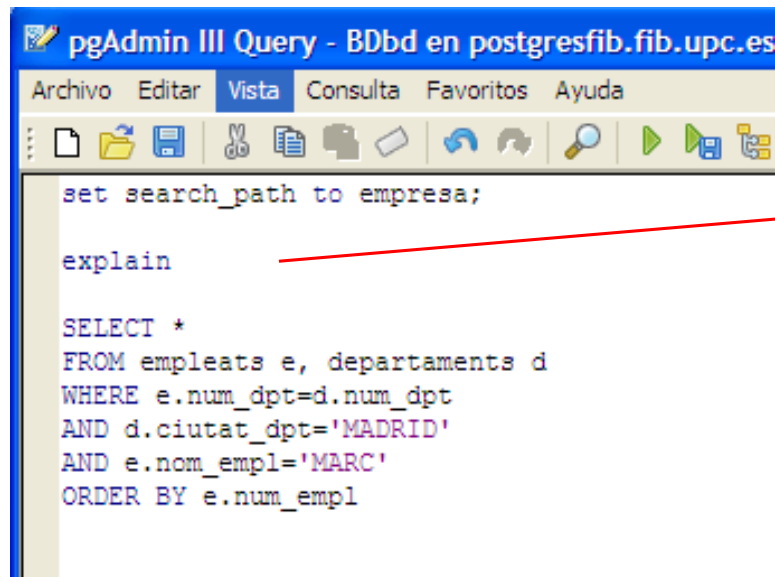
Declaratiu Que?

Procedimental Com?

- **Interés** de l'àlgebra relacional:

- ajuda a entendre quines funcionalitats de consulta ha de proporcionar un llenguatge relacional
- la versió estandard actual de SQL incorpora operacions de l'àlgebra relacional
- els SGBD processen i optimitzen les consultes basant-se en l'àlgebra relacional (recordar que l'àlgebra és procedimental i, per exemple, l'SQL és declaratiu)

SQL: QUÈ?



```
pgAdmin III Query - BDdb en postgresfib.fib.upc.es
Archivo  Editar  Vista  Consulta  Favoritos  Ayuda

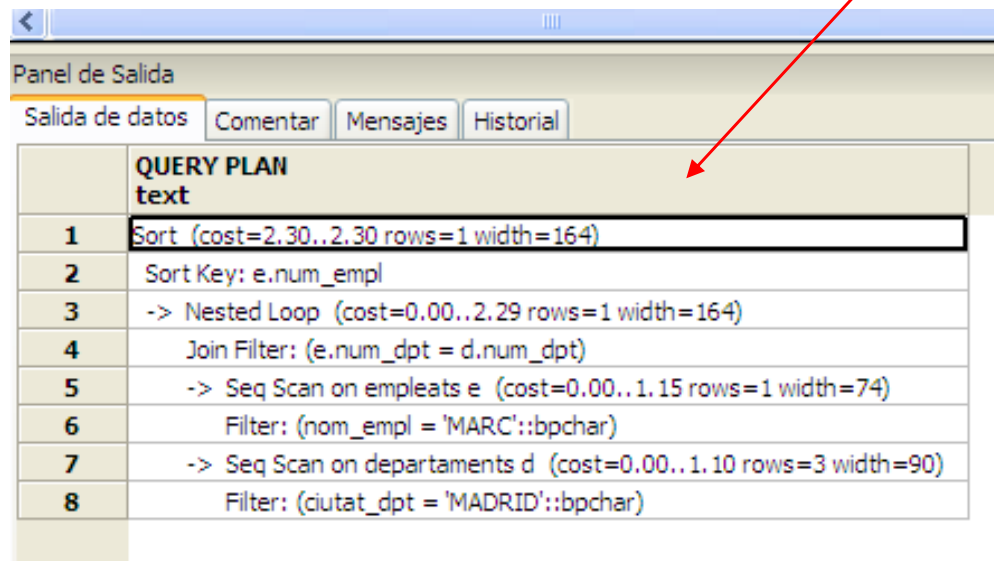
set search_path to empresa;

explain

SELECT *
FROM empleats e, departaments d
WHERE e.num_dpt=d.num_dpt
AND d.ciutat_dpt='MADRID'
AND e.nom_empl='MARC'
ORDER BY e.num_empl
```

Amb aquesta instrucció 'EXPLAIN' s'especifica al Postgres que "expliqui" COM executarà aquesta sentència

ÀLGEBRA: COM?



QUERY PLAN	
	text
1	Sort (cost=2.30..2.30 rows=1 width=164)
2	Sort Key: e.num_empl
3	-> Nested Loop (cost=0.00..2.29 rows=1 width=164)
4	Join Filter: (e.num_dpt = d.num_dpt)
5	-> Seq Scan on empleats e (cost=0.00..1.15 rows=1 width=74)
6	Filter: (nom_empl = 'MARC')::bpchar
7	-> Seq Scan on departaments d (cost=0.00..1.10 rows=3 width=90)
8	Filter: (ciutat_dpt = 'MADRID')::bpchar

El resultat conté entre d'altres coses (cost, amplada del resultat, algoritmes) l'expressió en àlgebra relacional que determina el millor ordre possible d'execució

4. Sort del resultat de 3

3. Join entre els resultats de 1 i 2

2. Selecció sobre empleats

1. Selecció sobre departaments

Operacions de l'àlgebra relacional

- **1a classif.:** Operacions conjuntistes

Unió
Intersecció
Diferència
Producte Cartesià

- **2a classif.:** Operacions primitives

Unió
Diferència
Producte Cartesià
Selecció
Projecció
Reanomenament

- **3a classif.:** Operacions binàries

Unió
Intersecció
Diferència
Producte Cartesià
Combinació (join)

- **Tancament relacional:** Tant els operands com el resultat d'una operació de l'àlgebra relacional són relacions

Ex: $T = R \cup S$

Operacions específicament relacionals

Selecció
Projecció
Combinació (join)
Reanomenament

Operacions no primitives

Intersecció
Combinació (join)

Operacions unàries

Selecció
Projecció
Reanomenament

Exemple

MODUL-CN(modul, sup-promig-de)

B6	10
B2	20

OFICINA(modul-de, num-de, superfície)

B6	25	10
B6	27	10
B2	25	15
B2	30	25

{modul-de} és una clau forana
que referencia MODUL-CN

PERSONAL-ADM(num-per, nom, cognom, modul, num)

100	Joan	Soler	B6	25
150	Clara	Bellsolà	B6	25

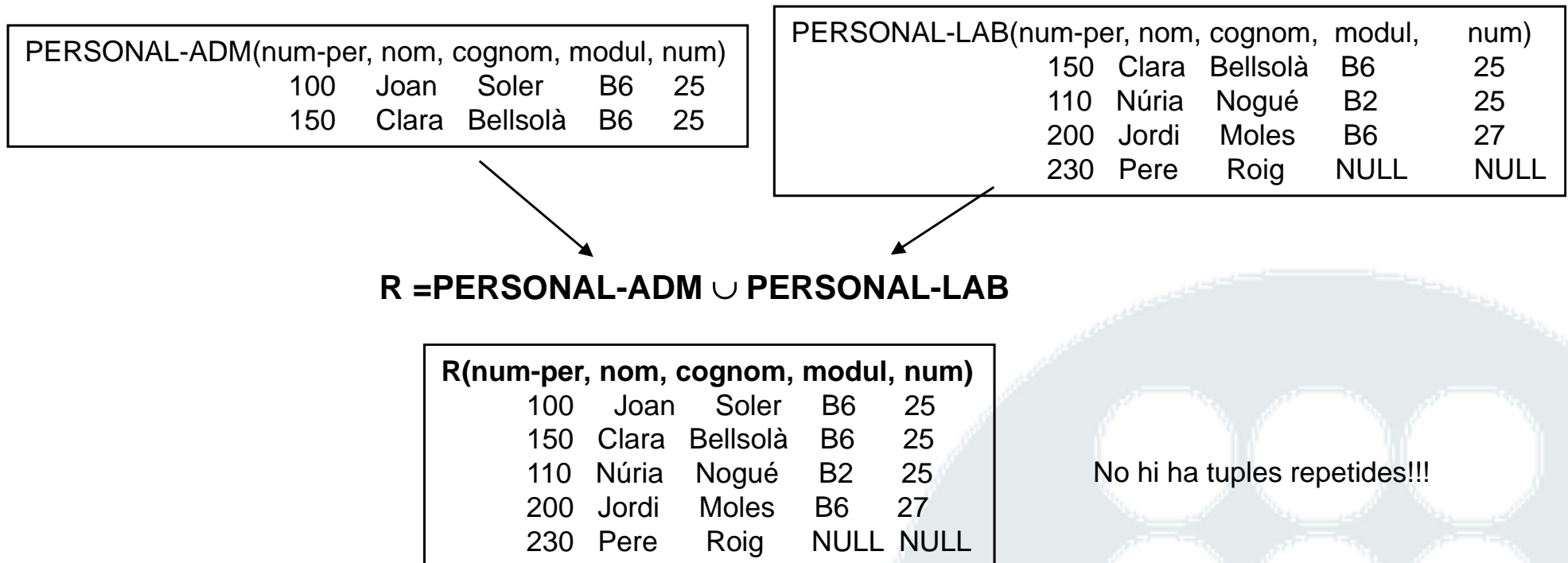
{modul, num} és una clau forana
que referencia OFICINA

PERSONAL-LAB(num-per, nom, cognom, modul, num)

150	Clara	Bellsolà	B6	25
110	Núria	Nogué	B2	25
200	Jordi	Moles	B6	27
230	Pere	Roig	NULL	NULL

{modul, num} és una clau forana
que referencia OFICINA

Unió



- Els atributs de l'**esquema** de la relació resultant de $T \cup S$ coincideixen amb els atributs de l'esquema de la relació T o de la relació S.
- L'**extensió** de la relació resultant de $T \cup S$ és el conjunt de tuples que pertanyen a l'extensió de T o que pertanyen a l'extensió de S o que pertanyen a l'extensió d'ambdues relacions
- Per fer la unió de dues relacions T i S cal que T i S siguin **relacions compatibles**.
- En cas de que els atributs de T i S no coincideixin cal reanomenar els atributs d'una de les dues relacions per tal de que siguin compatibles.

Relacions compatibles

- Algunes operacions de l'àlgebra relacional, com ara la unió, només té sentit que s'apliquin a relacions que siguin compatibles (que tinguin tuples "similars")

- Exemple: pot fer-se la unió

$\text{PERSONAL-ADM} \cup \text{PERSONAL-LAB}$

perquè les tuples de les dues relacions s'assemblen en canvi no té sentit fer la unió

$\text{PERSONAL-ADM} \cup \text{OFICINES}$

- Diem que **dues relacions T i S són compatibles** si:

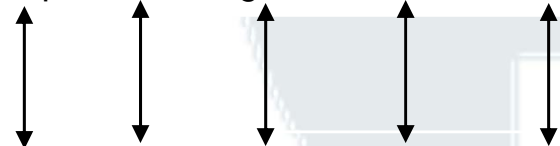
- tenen esquemes amb un conjunt d'atributs idèntic, i els dominis de cada parella d'atributs són els mateixos a T i a S.

Exemple:

PERSONAL-ADM i PERSONAL-LAB són clarament compatibles:

PERSONAL-ADM(num-per, nom, cognom, modul, num)

PERSONAL-LAB(num-per, nom, cognom, modul, num)



Reanomenament

PERSONAL-DOC(num-per, nom, cognom, modul-de, num-de)					
400	Jaume	Cases	Omega	119	
500	Pau	Pou	B6	123	

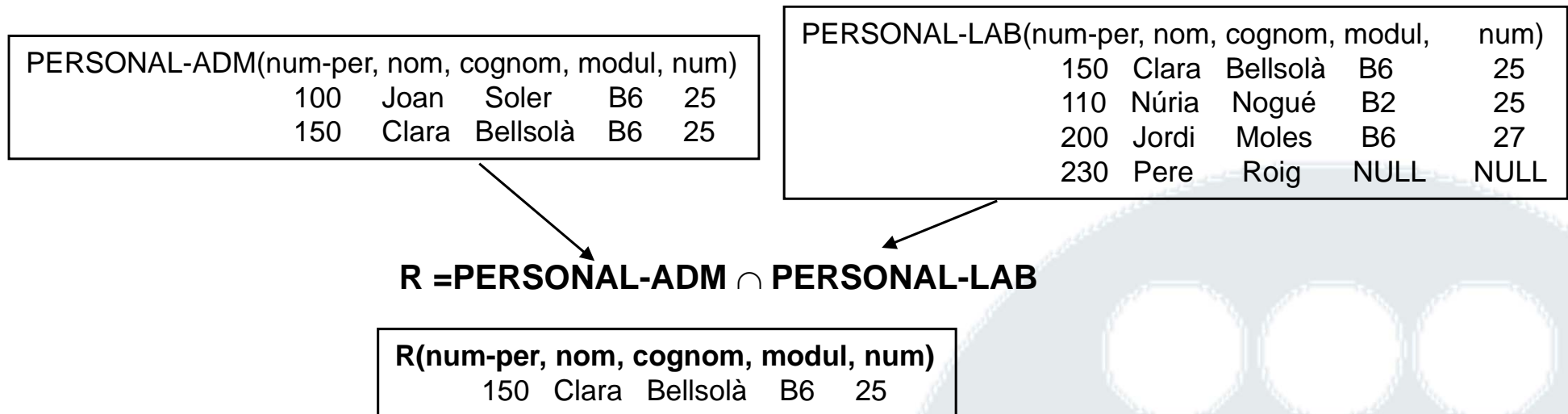


R = PERSONAL-DOC {modul-de -> modul, num-de -> num}

R(num-per, nom, cognom, modul, num)					
400	Jaume	Cases	Omega	119	
500	Pau	Pou	B6	123	

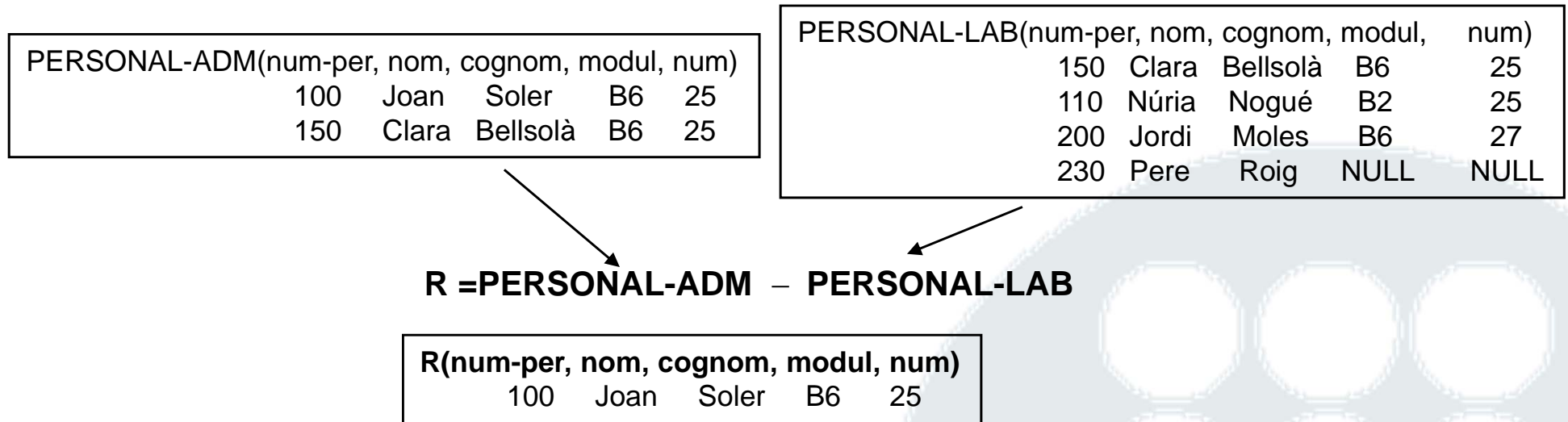
- L'**esquema** de la relació resultant és el mateix, exceptuant el canvi de nom dels atributs que han estat reanomenats.
- L'**extensió** de la relació resultant no canvia.

Intersecció



- Els atributs de l'**esquema** de la relació resultant de $T \cap S$ coincideixen amb els atributs de l'esquema de la relació T o la relació S.
- L'**extensió** de la relació resultant de $T \cap S$ és el conjunt de tuples que pertanyen a l'extensió d'ambdues relacions
- Per fer la intersecció de dues relacions T i S cal que T i S siguin **relacions compatibles**.
- En cas de que els atributs de T i S no coincideixin cal reanomenar els atributs d'una de les dues relacions per tal de que siguin compatibles.

Diferència



- Els atributs de l'**esquema** de la relació resultant de $T - S$ coincideixen amb els atributs de l'esquema de la relació T o la relació S.
- L'**extensió** de la relació resultant de $T - S$ és el conjunt de tuples que pertanyen a l'extensió de T però no a la de S.
- Per fer la diferència de dues relacions T i S cal que T i S siguin **relacions compatibles**.
- En cas de que els atributs de T i S no coincideixin cal reanomenar els atributs d'una de les dues relacions per tal de que siguin compatibles.

Producte Cartesià

MODUL-CN(modul, sup-promig-of)		
B6	10	
B2	20	

OFICINA(modul-de, num-de, superfície)		
B6	25	10
B6	27	10
B2	25	15
B2	30	25

$$R = \text{MODUL-CN} \times \text{OFICINA}$$

R(modul, sup-promig-of, modul-de, num-de, superfície)				
B6	10	B6	25	10
B6	10	B6	27	10
B6	10	B2	25	15
B6	10	B2	30	25
B2	20	B6	25	10
B2	20	B6	27	10
B2	20	B2	25	15
B2	20	B2	30	25

- Els atributs de l'**esquema** de la relació resultant de $T \times S$ són tots els atributs de T i tots els atributs de S.
- Si T i S tenen algun nom d'atribut idèntic, s'haurà de fer prèviament una operació de reanomenament d'una de les dues relacions per eliminar l'ambigüitat.
- L'**extensió** de la relació resultant de $T \times S$ és el conjunt de totes les tuples de la forma $\langle v_1, v_2, \dots, v_n, w_1, w_2, \dots, w_m \rangle$ on es compleix que $\langle v_1, v_2, \dots, v_n \rangle$ pertany a l'extensió de T i que $\langle w_1, w_2, \dots, w_m \rangle$ pertany a l'extensió de S.
- Per fer el producte cartesià de dues relacions T i S **no cal** que T i S siguin **relacions compatibles**

Selecció

OFICINA(modul-de, num-de, superfície)		
B6	25	10
B6	27	10
B2	25	15
B2	30	25



R =OFICINA(modul-de='B2' AND superfície>16)

R(modul-de, num-de, superfície)		
B2	30	25

- **T(C)** indica la selecció de T amb la condició C, essent C una condició de selecció
- La condició C està formada per una o més comparacions de la forma:

$$A_i \theta V_j \text{ o bé } A_i \theta A_j$$

on A_i i A_j són atributs de la relació T, V_j és un valor constant, i θ és un operador de comparació ($=$, $<>$, $<$, $<=$, $>$, $>=$). Les comparacions han d'estar relacionades entre elles per un dels operadors lògics AND (\wedge), OR (\vee).

- Els atributs de l'**esquema** de la relació resultant de T(C), coincideixen amb els atributs l'esquema de la relació T
- L'**extensió** de la relació resultant de T(C) és el conjunt de tuples que pertanyen a l'extensió de T i que satisfan la condició de selecció C.

Projecció

PERSONAL-ADM(num-per, nom, cognom, modul, num)					
100	Joan	Soler	B6	25	
150	Clara	Bellsolà	B6	25	



$R = \text{PERSONAL-ADM}[\text{modul}, \text{num}]$

R(modul, num)	
B6	25

En àlgebra relacional,
NO HI HA TUPLES REPETIDES
 perquè el resultat de les
 operacions són conjunts.

- $T[A_i, A_j, \dots, A_k]$ indica la projecció de T sobre $\{A_i, A_j, \dots, A_k\}$, essent $\{A_i, A_j, \dots, A_k\}$ un subconjunt dels atributs de l'esquema de la relació T.
- Els atributs de l'**esquema** de la relació resultant de $T[A_i, A_j, \dots, A_k]$, són els atributs $\{A_i, A_j, \dots, A_k\}$
- L'**extensió** de la relació resultant de $T[A_i, A_j, \dots, A_k]$ és el conjunt de totes les tuples de la forma $\langle t.A_i, t.A_j, \dots, t.A_k \rangle$ on es compleix que t és una tupla de l'extensió de T i on $t.A_p$ denota el valor per l'atribut A_p de la tupla t.

Combinació (join)

MODUL-CN(modul, sup-promig-of)		
B6	10	
B2	20	

OFICINA(modul-de, num-de, superfície)		
B6	25	10
B6	27	10
B2	25	15
B2	30	25

R = MODUL-CN[modul=modul-de]OFICINA

R(modul, sup-promig-of, modul-de, num-de, superfície)				
B6	10	B6	25	10
B6	10	B6	27	10
B2	20	B2	25	15
B2	20	B2	30	25

R = MODUL-CN[modul=modul-de, sup-promig-of ≤ superfície]OFICINA

R(modul, sup-promig-of, modul-de, num-de, superfície)				
B6	10	B6	25	10
B6	10	B6	27	10
B2	20	B2	30	25

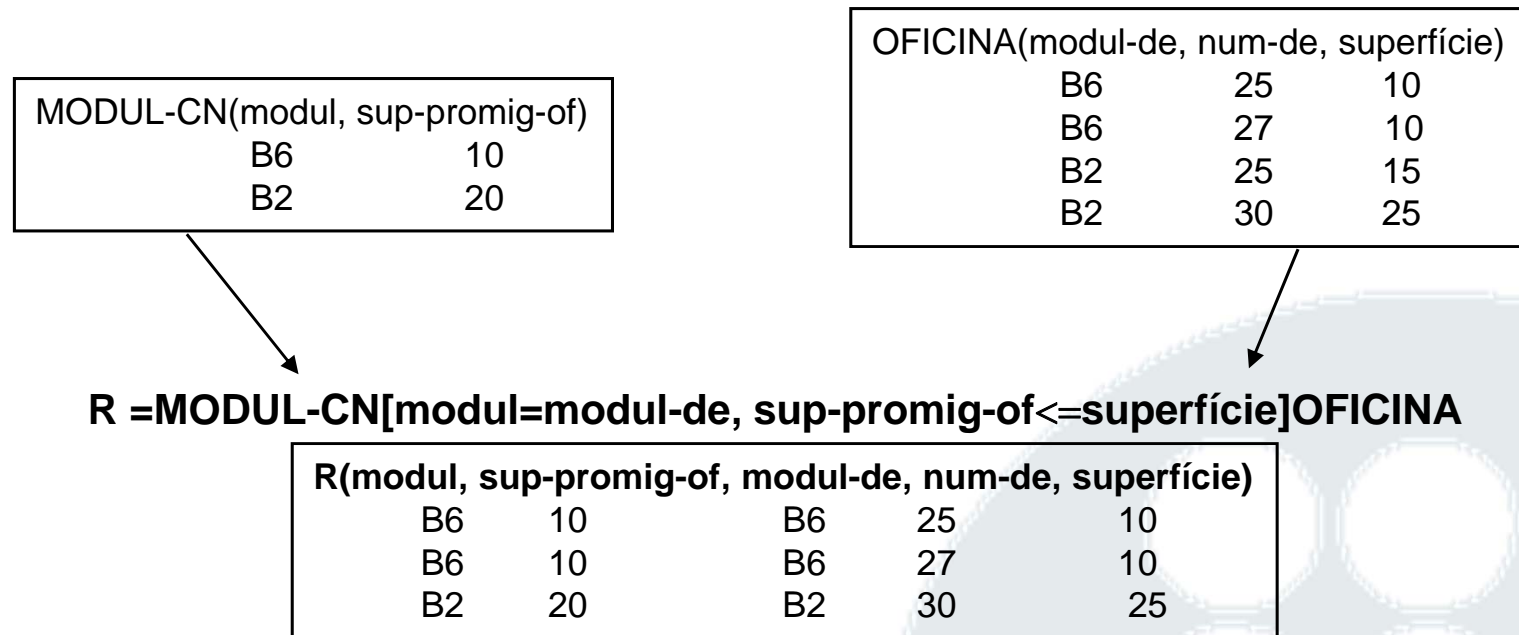
- **T[B]S** indica la combinació de T i S amb la condició B
- La **condició B** d'una combinació **T [B] S** està formada per una o més comparacions de la forma:

$$A_i \theta A_j$$

on A_i és un atribut de la relació T, A_j és un atribut de la relació S, θ és un operador de comparació ($=, \neq, <, \leq, >, \geq$) i es compleix que A_i i A_j tenen el mateix domini.

- Les comparacions d'una condició de combinació se separen per comes.

Combinació (join)



- Els atributs de l'**esquema** de la relació resultant de $T[B]S$ són tots els atributs de T i tots els atributs de S .
- Si T i S tenen algun nom d'atribut idèntic, s'haurà de fer prèviament una operació de reanomenament d'una de les dues relacions per eliminar l'ambigüitat.
- L'**extensió** de la relació resultant de $T[B]S$ és el conjunt de tuples que pertanyen a l'extensió del producte cartesià $T \times S$ i que satisfan totes les comparacions que formen la condició de combinació B .

Combinació (join): Tipus de “joins”

- **“θ-join”**: La “join” s’anomena també “θ-join”
- **“Equi-join”**: Cas particular de “join” en què totes les comparacions de la condició tenen l’operador ‘=’.

PERSONAL-ADM(num-per, nom, cognom, modul, num)					
100	Joan	Soler	B6	25	
150	Clara	Bellsolà	B6	25	

OFICINA(modul-de, num-de, superfície)		
B6	25	10
B6	27	10
B2	25	15
B2	30	25

R = PERSONAL-ADM[modul=modul-de, num=num-de]OFICINA

R(num-per, nom, cognom, modul, num, modul-de, num-de, superfície)							
100	Joan	Soler	B6	25	B6	25	10
150	Clara	Bellsolà	B6	25	B6	25	10

- **“Natural join”**: Variant de la “equi-join” en la qual s’eliminen els atributs superflus. Es denota mitjançant un *. La diferència amb la “equi-join” és en l’esquema de la relació resultant, ja que no hi apareix el segon atribut de la comparació.

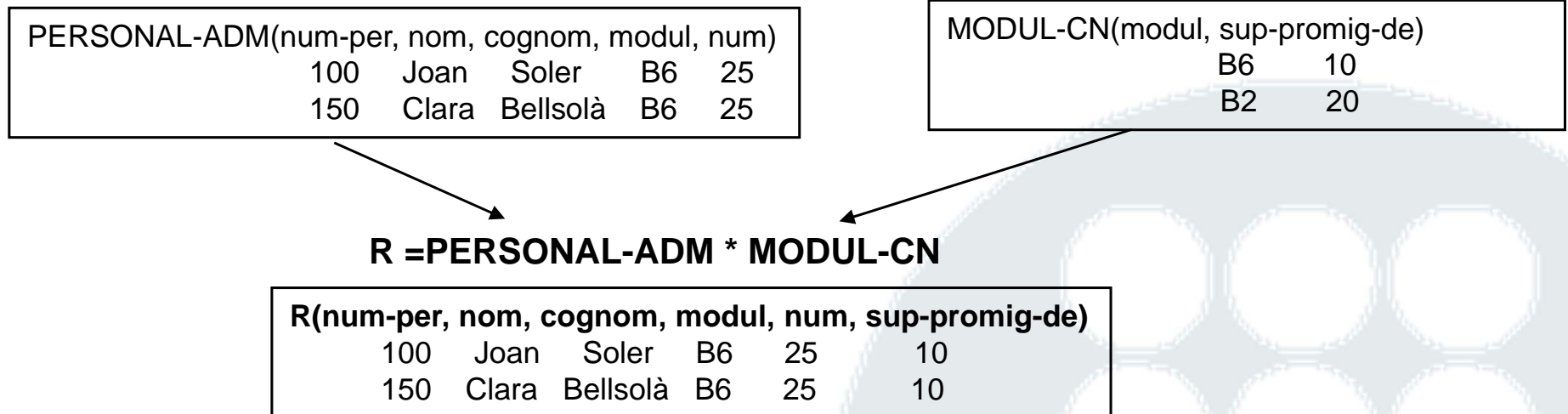
PERSONAL-ADM(num-per, nom, cognom, modul, num)					
100	Joan	Soler	B6	25	
150	Clara	Bellsolà	B6	25	

OFICINA(modul-de, num-de, superfície)		
B6	25	10
B6	27	10
B2	25	15
B2	30	25

R = PERSONAL-ADM[modul*modul-de, num*num-de]OFICINA

R(num-per, nom, cognom, modul, num, superfície)						
100	Joan	Soler	B6	25	10	
150	Clara	Bellsolà	B6	25	10	

Combinació (join): “Natural join” implícita



- La “**natural join**” **ímplicita**: Variant de la “natural-join” en la qual no s’especifica la condició de combinació i aleshores s’assumeix per defecte que la condició de combinació correspon a la d’una “natural join” on s’igualen tots els parells d’atributs que tenen el mateix nom a les dues relacions.
- **T * S** denota la “natural join” ímplicita de T i S.

Seqüències d'operacions de l'àlgebra relacional

Exemple: Obtenir les oficines (modul i número) dels moduls que tenen una superfície promig més gran que 15.

MODUL-CN(modul, sup-promig-of)		
B6		10
B2		20

OFICINA(modul-de, num-de, superfície)		
B6	25	10
B6	27	10
B2	25	15
B2	30	25

A = MODUL-CN(sup-promig-of > 15)

B = A{modul -> modul-de}

C = OFICINA * B

R = C[modul-de, num-de]

R(modul-de, num-de)	
B2	25
B2	30

- Les **consultes** a una BD relacional es poden expressar en termes de **seqüències d'operacions** de l'àlgebra relacional.
- Les seqüències d'operacions ens permeten definir una relació que conté precisament allò que es desitja consultar.