

# Bases de Dados

## Parte IV

### Álgebra e Cálculo Relacional

### Álgebra Relacional

- Conjunto básico de operações que nos permitem manipular relações no modelo relacional.
- As operações da álgebra relacional produzem novas relações, ou seja, a aplicação de uma operação da álgebra relacional tem sempre como resultado uma **nova relação**.
- As relações obtidas por utilização das operações da álgebra relacional podem ser igualmente utilizadas em outras operações da álgebra.
- Uma sequência de operações da álgebra relacional forma uma expressão cujo resultado é uma relação que representa o resultado de uma **consulta à base de dados**.
- A álgebra relacional é utilizada principalmente como formalismo para implementar e otimizar consultas no modelo relacional.
- A linguagem SQL incorpora alguns dos conceitos da álgebra relacional.

## Operações da Álgebra Relacional

- Operações específicas de BD
  - Selecção  $\sigma$
  - Projecção  $\pi$
  - Renomeação  $\rho$
  - Junção  $\bowtie$
  - Agregação  $\mathcal{F}$
- Operações sobre conjuntos
  - Reunião  $\cup$
  - Intersecção  $\cap$
  - Diferença  $-$
  - Produto cartesiano  $\times$
  - Divisão  $\div$

## Operação de Selecção

- Permite seleccionar um subconjunto de tuplos de uma relação a partir de uma condição sobre os atributos.
- A operação de selecção é representada pela expressão

$$\sigma_{\text{COND}}(R)$$

em que  $\sigma$  é o operador de selecção e COND é a condição sobre os atributos da relação R.

- A condição COND pode ser da forma

$$A_i \theta A_j \quad \text{ou} \quad A_i \theta \text{VAL}$$

em que  $A_i$  e  $A_j$  são atributos,  $\theta$  é um operador de comparação  $\{=, <, >, \leq, \geq, \neq\}$  e  $\text{VAL} \in \text{dom}(A_i)$ .

- A condição COND também pode ser composta por várias cláusulas ligadas pelos operadores lógicos AND, OR ou NOT.

## Operação de Seleção

EMPREGADO	NomeP	NomeF	NumBI	...	Salário	NumDep
	João	Santos	798764544	...	2500	4
	Inês	Pereira	345673451	...	2000	1
	Rui	Silva	487563546	...	1500	2
	Ana	Feio	342342324	...	3000	4

- Obtenha os empregados que trabalham no departamento 4 e cujo salário é superior a 2000 euros.

$$\sigma_{\text{NumDep} = 4 \text{ AND Salário} > 2000}(\text{EMPREGADO})$$

NomeP	NomeF	NumBI	...	Salário	NumDep
João	Santos	798764544	...	2500	4
Ana	Feio	342342324	...	3000	4

## Operação de Seleção

- A relação resultante da operação de selecção

$$\sigma_{\text{COND}}(\text{R})$$

tem os mesmos atributos da relação R.

- $\sigma$  é um operador comutativo:

$$\sigma_{\text{COND1}}(\sigma_{\text{COND2}}(\text{R})) = \sigma_{\text{COND2}}(\sigma_{\text{COND1}}(\text{R}))$$

- Uma sequência de operações de selecção pode ser substituída por uma única operação de selecção com a conjunção de todas as condições:

$$\sigma_{\text{COND1}}(\sigma_{\text{COND2}}(\sigma_{\text{COND3}}(\text{R}))) = \sigma_{\text{COND1 AND COND2 AND COND3}}(\text{R})$$

## Operação de Projecção

- Permite obter uma nova relação com apenas alguns atributos da relação original.
- A operação de projecção é representada pela expressão

$$\pi_{\text{ATRIBS}}(\text{R})$$

em que  $\pi$  é o operador de projecção e ATRIBS é a lista de atributos a seleccionar da relação R.

- A relação resultante da operação de projecção tem apenas os atributos definidos em ATRIBS na mesma ordem em que estes aparecem indicados na lista.

## Operação de Projecção

EMPREGADO	NomeP	NomeF	NumBI	...	Salário	NumDep
	João	Santos	798764544	...	2500	4
	Inês	Pereira	345673451	...	2000	1
	Rui	Silva	487563546	...	1500	2
	Ana	Feio	342342324	...	3000	4

- Obtenha o número do BI, primeiro e último nome de todos os empregados.

$$\pi_{\text{NumBI, NomeP, NomeF}}(\text{EMPREGADO})$$

NumBI	NomeP	NomeF
798764544	João	Santos
345673451	Inês	Pereira
487563546	Rui	Silva
342342324	Ana	Feio

## Operação de Projecção

- Para que o resultado seja uma relação válida, a operação de projecção **remove tuplos repetidos** (pode ser necessário quando a lista não inclui a chave de R).
- Uma sequência de operações de projecção pode ser substituída por uma única operação de projecção:

$$\pi_{\text{ATRIBS1}}(\pi_{\text{ATRIBS2}}(R)) = \pi_{\text{ATRIBS1}}(R)$$

se  $\text{ATRIBS1} \subseteq \text{ATRIBS2}$  pois caso contrário é uma expressão inválida.

## Operação de Projecção

EMPREGADO	NomeP	NomeF	NumBI	...	Salário	NumDep
	João	Santos	798764544	...	2500	4
	Inês	Pereira	345673451	...	2000	1
	Rui	Silva	487563546	...	1500	2
	Ana	Feio	342342324	...	3000	4

- Obtenha o número dos departamentos onde trabalha pelo menos um empregado.

$$\pi_{\text{NumDep}}(\text{EMPREGADO})$$

NumDep
4
1
2

## Resultados Intermediários

EMPREGADO	NomeP	NomeF	NumBI	...	Salário	NumDep
	João	Santos	798764544	...	2500	4
	Inês	Pereira	345673451	...	2000	1
	Rui	Silva	487563546	...	1500	2
	Ana	Feio	342342324	...	3000	4

- Obtenha o número do BI, primeiro e último nome dos empregados que trabalham no departamento 4 e cujo salário é superior a 2000 euros.

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

NumBI	NomeP	NomeF
798764544	João	Santos
342342324	Ana	Feio

## Resultados Intermediários

- A consulta anterior pode ser escrita criando resultados intermediários sobre os quais vamos aplicando as operações restantes.

$$\text{DEP4\_SAL2000} \leftarrow \sigma_{\text{NumDep} = 4 \text{ AND Salário} > 2000}(\text{EMPREGADO})$$

$$\text{RESULT} \leftarrow \pi_{\text{NumBI, NomeP, NomeF}}(\text{DEP4\_SAL2000})$$

DEP4_SAL2000	NomeP	NomeF	NumBI	...	Salário	NumDep
	João	Santos	798764544	...	2500	4
	Ana	Feio	342342324	...	3000	4

RESULT	NumBI	NomeP	NomeF
	798764544	João	Santos
	342342324	Ana	Feio

## Operação de Renomeação

- Permite obter uma nova relação com o nome da relação e/ou o nome dos atributos renomeados.
- A operação de renomeação é representada pelas expressões

$$\rho_S(R) \quad \text{ou} \quad \rho_{(B_1, B_2, \dots, B_n)}(R) \quad \text{ou} \quad \rho_{S(B_1, B_2, \dots, B_n)}(R)$$

em que  $\rho$  é o operador de renomeação,  $S$  é o novo nome da relação e  $B_1, B_2, \dots, B_n$  são os novos nomes dos atributos.

- A consulta anterior pode ser escrita usando o operador de renomeação.

$$\rho_{\text{DEP4\_SAL2000}}(\sigma_{\text{NumDep} = 4 \text{ AND Salário} > 2000}(\text{EMPREGADO}))$$

$$\rho_{\text{RESULT}}(\text{BI, Nome, Apelido})(\pi_{\text{NumBI, NomeP, NomeF}}(\text{DEP4\_SAL2000}))$$

RESULT	BI	Nome	Apelido
	798764544	João	Santos
	342342324	Ana	Feio

## Operações de Reunião, Intersecção e Diferença

- As operações de reunião, intersecção e diferença são operações binárias sobre relações que correspondem às operações matemáticas básicas sobre conjuntos.
- As operações de reunião, intersecção e diferença são representadas pelas expressões

$$R \cup S \quad \text{e} \quad R \cap S \quad \text{e} \quad R - S$$

em que  $\cup, \cap$  e  $-$  são os operadores de reunião, intersecção e diferença e  $R$  e  $S$  são duas **relações compatíveis para a reunião**.

- Duas relações  $R(A_1, A_2, \dots, A_n)$  e  $S(B_1, B_2, \dots, B_n)$  dizem-se compatíveis para a reunião se tiverem o mesmo grau  $n$  e se  $\text{dom}(A_i) = \text{dom}(B_i)$  para  $1 \leq i \leq n$ .
- Por convenção, a relação resultante das operações  $R \cup S$ ,  $R \cap S$  e  $R - S$  tem os nomes dos atributos da relação  $R$ .

## Operações de Reunião, Intersecção e Diferença

- O resultado da operação  $R \cup S$  é a relação que inclui todos os tuplos que estão em R, em S ou em ambos. **Tuplos repetidos são removidos.**
- O resultado da operação  $R \cap S$  é a relação que inclui todos os tuplos que estão em R e em S.
- O resultado da operação  $R - S$  é a relação que inclui todos os tuplos que estão em R mas não em S.
- $\cup$  e  $\cap$  são operadores comutativos e associativos:
 
$$R \cup S = S \cup R$$

$$R \cap S = S \cap R$$

$$R \cup (S \cup T) = (R \cup S) \cup T$$

$$R \cap (S \cap T) = (S \cap R) \cap T$$
- O operador  $-$  não é comutativo nem associativo, pois em geral:
 
$$R - S \neq S - R$$

## Operação de Reunião

EMPREGADO	NomeP	NomeF	NumBI	...	Salário	SuperBI	NumDep
	João	Santos	798764544	...	2500	487563546	4
	Inês	Pereira	345673451	...	2000	487563546	1
	Rui	Silva	487563546	...	1500	123456789	2
	Ana	Feio	342342324	...	3000	798764544	4

- Obtenha o número do BI dos empregados que trabalham no departamento 4 ou que supervisionam um empregado que trabalha no departamento 4.

$$\text{EMP\_DEP4} \leftarrow \sigma_{\text{NumDep} = 4}(\text{EMPREGADO})$$

$$\text{RESULT1} \leftarrow \pi_{\text{NumBI}}(\text{EMP\_DEP4})$$

$$\text{RESULT2} \leftarrow \pi_{\text{SuperBI}}(\text{EMP\_DEP4})$$

$$\text{RESULT} \leftarrow \text{RESULT1} \cup \text{RESULT2}$$

RESULT	NumBI
	798764544
	342342324
	487563546

### Operação de Intersecção

EMPREGADO	NomeP	NomeF	NumBI	...	Salário	SuperBI	NumDep
	João	Santos	798764544	...	2500	487563546	4
	Inês	Pereira	345673451	...	2000	487563546	1
	Rui	Silva	487563546	...	1500	123456789	2
	Ana	Feio	342342324	...	3000	798764544	4

- Obtenha o número do BI dos empregados que trabalham no departamento 4 e que supervisionam um empregado que trabalha no departamento 4.

$$\text{EMP\_DEP4} \leftarrow \sigma_{\text{NumDep} = 4}(\text{EMPREGADO})$$

$$\text{RESULT1} \leftarrow \pi_{\text{NumBI}}(\text{EMP\_DEP4})$$

$$\text{RESULT2} \leftarrow \pi_{\text{SuperBI}}(\text{EMP\_DEP4})$$

$$\text{RESULT} \leftarrow \text{RESULT1} \cap \text{RESULT2}$$

RESULT	NumBI
	798764544

### Operação de Diferença

EMPREGADO	NomeP	NomeF	NumBI	...	Salário	SuperBI	NumDep
	João	Santos	798764544	...	2500	487563546	4
	Inês	Pereira	345673451	...	2000	487563546	1
	Rui	Silva	487563546	...	1500	123456789	2
	Ana	Feio	342342324	...	3000	798764544	4

- Obtenha o número do BI dos empregados que trabalham no departamento 4 e que não supervisionam um empregado que trabalha no departamento 4.

$$\text{EMP\_DEP4} \leftarrow \sigma_{\text{NumDep} = 4}(\text{EMPREGADO})$$

$$\text{RESULT1} \leftarrow \pi_{\text{NumBI}}(\text{EMP\_DEP4})$$

$$\text{RESULT2} \leftarrow \pi_{\text{SuperBI}}(\text{EMP\_DEP4})$$

$$\text{RESULT} \leftarrow \text{RESULT1} - \text{RESULT2}$$

RESULT	NumBI
	342342324

## Produto Cartesiano

- Permite combinar numa nova relação todos os tuplos de duas relações.
- O produto cartesiano é representado pela expressão
 
$$R \times S$$
 em que  $\times$  é o operador de produto cartesiano e R e S são duas relações quaisquer.
- O produto cartesiano entre  $R(A_1, A_2, \dots, A_n)$  e  $S(B_1, B_2, \dots, B_m)$  tem como resultado a relação  $T(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$  onde existe um tuplo para cada combinação possível de tuplos de R com tuplos de S.
- Se R tiver NR tuplos e S tiver NS tuplos, então  $R \times S$  tem  $NR * NS$  tuplos.

## Produto Cartesiano

R	A1	A2
	a1_val1	a2_val1
	a1_val2	a2_val2
	a1_val3	a2_val3

S	B1	B2	B3
	b1_val1	b2_val1	b3_val1
	b1_val2	b2_val2	b3_val2

R × S	A1	A2	B1	B2	B3
	a1_val1	a2_val1	b1_val1	b2_val1	b3_val1
	a1_val1	a2_val1	b1_val2	b2_val2	b3_val2
	a1_val2	a2_val2	b1_val1	b2_val1	b3_val1
	a1_val2	a2_val2	b1_val2	b2_val2	b3_val2
	a1_val3	a2_val3	b1_val1	b2_val1	b3_val1
	a1_val3	a2_val3	b1_val2	b2_val2	b3_val2

## Produto Cartesiano

- Obtenha o nome dos dependentes dos empregados do sexo feminino.

$$\text{EMP\_FEM} \leftarrow \sigma_{\text{Sexo} = 'F'}(\text{EMPREGADO})$$

$$\text{FEM\_DEPS} \leftarrow \sigma_{\text{NumBI} = \text{EmpBI}}(\text{EMP\_FEM} \times \text{DEPENDENTES})$$

$$\text{RESULT} \leftarrow \pi_{\text{Nome}}(\text{FEM\_DEPS})$$

Alternativa de resolução mais eficiente:

$$\text{EMP\_FEM} \leftarrow \pi_{\text{NumBI}}(\sigma_{\text{Sexo} = 'F'}(\text{EMPREGADO}))$$

$$\text{DEPS\_NOME} \leftarrow \pi_{\text{EmpBI, Nome}}(\text{DEPENDENTES})$$

$$\text{FEM\_DEPS} \leftarrow \sigma_{\text{NumBI} = \text{EmpBI}}(\text{EMP\_FEM} \times \text{DEPS\_NOME})$$

$$\text{RESULT} \leftarrow \pi_{\text{Nome}}(\text{FEM\_DEPS})$$

## Operação de Junção

- Permite combinar tuplos de duas relações que obedecem a uma condição de junção.
- A operação de junção (ou junção- $\theta$ ) é representada pela expressão

$$R \bowtie_{\text{COND}} S$$

em que  $\bowtie$  é o operador de junção, COND é a condição de junção e R e S são duas relações quaisquer.

- A condição COND é da forma

$$A_i \theta B_j$$

em que  $\theta$  é um operador de comparação  $\{=, <, >, \leq, \geq, \neq\}$ ,  $A_i$  é um atributo de R e  $B_j$  é um atributo de S.

- A condição COND também pode ser composta por várias cláusulas ligadas pelo operador lógico AND.

## Operação de Junção

- A operação de junção entre  $R(A_1, A_2, \dots, A_n)$  e  $S(B_1, B_2, \dots, B_m)$  tem como resultado a relação  $T(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$  onde existe um tuplo para cada combinação possível de tuplos de R com tuplos de S que satisfaz a condição de junção.

- A operação de junção equivale às operações de produto cartesiano e de selecção:

$$R \bowtie_{\text{COND}} S = \sigma_{\text{COND}}(R \times S)$$

- A operação de junção é muito utilizada em BD pois permite lidar com os relacionamentos entre relações.

## Operação de Junção

- Obtenha o nome dos dependentes dos empregados do sexo feminino.

$$\text{EMP\_FEM} \leftarrow \sigma_{\text{Sexo} = \text{'F'}}(\text{EMPREGADO})$$

$$\text{RESULT} \leftarrow \pi_{\text{Nome}}(\text{EMP\_FEM} \bowtie_{\text{NumBI} = \text{EmpBI}} \text{DEPENDENTES})$$

- Obtenha o nome e o endereço de todos os empregados que trabalham no departamento de Produção.

$$\text{EMP\_DEP} \leftarrow \text{EMPREGADO} \bowtie_{\text{NumDep} = \text{Num}} \text{DEPARTAMENTO}$$

$$\text{RESULT} \leftarrow \pi_{\text{NomeP}, \text{NomeF}, \text{Endereço}}(\sigma_{\text{Nome} = \text{'Produção'}}(\text{EMP\_DEP}))$$

- Para todos os projectos localizados no Porto, obtenha o nome do projecto e o último nome do respectivo gerente.

$$\text{PROJ\_PORTO} \leftarrow \sigma_{\text{PLocal} = \text{'Porto'}}(\rho_{(\text{PNome}, \text{PNum}, \text{PLocal}, \text{PDep})}(\text{PROJECTO}))$$

$$\text{PORTO\_DEP} \leftarrow \text{PROJ\_PORTO} \bowtie_{\text{PDep} = \text{Num}} \text{DEPARTAMENTO}$$

$$\text{RESULT} \leftarrow \pi_{\text{PNome}, \text{NomeF}}(\text{PORTO\_DEP} \bowtie_{\text{GerenteBI} = \text{NumBI}} \text{EMPREGADO})$$

## Operação de Equi-Junção

- Operação de junção em que se utiliza apenas o operador de comparação '=' na condição de junção.
- A relação resultante de uma operação de equi-junção tem pelo menos um par de atributos em duplicado (com os mesmos valores em todos os tuplos).
- Os atributos em duplicado são os atributos envolvidos na condição de junção. Por exemplo, na relação DEP\_GER, GerenteBI e NumBI têm os mesmos valores em todos os tuplos:

$$\text{DEP\_GER} \leftarrow \text{DEPARTAMENTO} \bowtie_{\text{GerenteBI} = \text{NumBI}} \text{EMPREGADO}$$

DEP_GER	Nome	Num	GerenteBI	GerenteData	NomeP	NomeF	NumBI	...
	Vendas	4	798764544	01-01-2000	João	Santos	798764544	...
	Produção	1	345673451	15-10-2003	Inês	Pereira	345673451	...
	Pós-Venda	2	487563546	01-06-1995	Rui	Silva	487563546	...

## Operação de Junção Natural

- Operação de equi-junção em que para cada par de atributos em duplicado um dos atributos é removido.
- A operação de junção natural (ou junção-\*) é representada pela expressão  $R * S$  em que \* é o operador de junção natural e R e S são duas relações com pelo menos um par de atributos com o mesmo nome (atributos de junção).
- Os atributos de junção têm de ter o mesmo nome, pelo que pode ser necessário fazer renomeação de atributos.

$$\rho_{(\text{Nome}, \text{Num}, \text{NumBI}, \text{GerenteData})}(\text{DEPARTAMENTO}) * \text{EMPREGADO}$$

Nome	Num	NumBI	GerenteData	NomeP	NomeF	...
Vendas	4	798764544	01-01-2000	João	Santos	...
Produção	1	345673451	15-10-2003	Inês	Pereira	...
Pós-Venda	2	487563546	01-06-1995	Rui	Silva	...

## Operações de Junção Interna e de Junção Externa

- Nas operações de junção anteriores, todos os tuplos que não satisfazem a condição de junção são removidos da relação resultante. As operações deste tipo são designadas por operações de **junção interna**.
- Um outro conjunto de operações permite manter parte ou a totalidade dos tuplos das relações a combinar mesmo que estes não satisfaçam a condição de junção. As operações deste tipo são designadas por operações de **junção externa**.
- Existem três tipos de operações de junção externa:
  - Junção externa à esquerda  $\bowtie$
  - Junção externa à direita  $\bowtie$
  - Junção externa completa  $\bowtie$

## Operações de Junção Externa

- Junção externa à esquerda ( $R \bowtie_{COND} S$ )
  - Mantém todos os tuplos da relação à esquerda R, e para os tuplos que não satisfazem a condição de junção preenche os atributos da relação à direita S com valores NULL.
- Junção externa à direita ( $R \bowtie_{COND} S$ )
  - Mantém todos os tuplos da relação à direita S, e para os tuplos que não satisfazem a condição de junção preenche os atributos da relação à esquerda R com valores NULL.
- Junção externa completa ( $R \bowtie_{COND} S$ )
  - Mantém todos os tuplos de ambas as relações e para os tuplos que não satisfazem a condição de junção preenche os atributos da relação combinada com valores NULL.

## Operações de Junção Externa

- Obtenha o número do BI dos empregados que não têm dependentes.

$$\text{EMP\_DEPS} \leftarrow \text{EMPREGADO} \bowtie_{\text{NumBI} = \text{EmpBI}} \text{DEPENDENTES}$$

$$\text{RESULT} \leftarrow \pi_{\text{NumBI}}(\sigma_{\text{EmpBI} = \text{NULL}}(\text{EMP\_DEPS}))$$

EMP_DEPS	NomeP	NomeF	NumBI	...	EmpBI	Nome	DataNasc	GrauParentesco
	João	Santos	798764544	...	798764544	Joana	23-05-1970	Esposa
	Inês	Pereira	345673451	...	NULL	NULL	NULL	NULL
	Rui	Silva	487563546	...	487563546	Maria	18-07-1975	Esposa
	Rui	Silva	487563546	...	487563546	Carlos	03-09-1999	Filho
	Ana	Feio	342342324	...	342342324	Pedro	22-11-1983	Marido

RESULT	NumBI
	345673451

## Operação de Divisão

- Permite obter os valores de uma relação que estão combinados com todos os tuplos de outra relação.
- A operação de divisão é representada pela expressão

$$R \div S$$

em que  $\div$  é o operador de divisão e R e S são duas relações em que os atributos de S são um subconjunto dos atributos de R.

- O resultado da operação  $R(Z) \div S(X)$  é a relação  $T(Y)$ , com  $Y = Z - X$ , que inclui todos os tuplos t para os quais existe um subconjunto R' de R tal que  $\pi_Y(R') = t$  e  $\pi_X(R') = S$ .
- A operação de divisão pode ser expressa utilizando os operadores  $\pi$ ,  $\times$  e  $-$ :

$$R(Z) \div S(X) = \pi_Y(R) - \pi_Y((S \times \pi_Y(R)) - R), \text{ com } Y = Z - X.$$

### Operação de Divisão

R	A	B
	a1	b1
	a2	b1
	a3	b1
	a4	b1
	a1	b2
	a3	b2
	a2	b3
	a3	b3
	a4	b3
	a1	b4
	a2	b4
	a3	b4

S	A
	a1
	a2
	a3

R ÷ S	B
	b1
	b4

### Operação de Divisão

- Obtenha o nome dos empregados que trabalham em TODOS os projectos nos quais o Rui Silva também trabalha.

$$\text{EMP\_SILVA} \leftarrow \sigma_{\text{NomeP} = \text{'Rui'} \text{ AND NomeF} = \text{'Silva'}}(\text{EMPREGADO})$$

$$\text{SILVA\_PROJ} \leftarrow \pi_{\text{NumProj}}(\text{TRABALHA\_EM} \bowtie_{\text{EmpBI} = \text{NumBI}} \text{EMP\_SILVA})$$

$$\text{BI\_PROJ} \leftarrow \pi_{\text{EmpBI}, \text{NumProj}}(\text{TRABALHA\_EM})$$

$$\text{RESULT\_BI} \leftarrow \rho_{(\text{NumBI})}(\text{BI\_PROJ} \div \text{SILVA\_PROJ})$$

$$\text{RESULT} \leftarrow \pi_{\text{NomeP}, \text{NomeF}}(\text{RESULT\_BI} * \text{EMPREGADO})$$

## Operação de Agregação

- Permite agrupar tuplos e sumariar informação a partir de funções de agregação.
- A operação de agregação é representada pela expressão

$$\text{ATRIBS } \mathcal{F} \text{ FUNS}(\text{R})$$

em que  $\mathcal{F}$  é o operador de agregação, ATRIBS é a lista de atributos de R a agrupar e FUNS é a lista das funções de agregação.

- A lista FUNS é da forma

$$\theta A_i$$

em que  $\theta$  é uma função de agregação {SUM, AVERAGE, MAXIMUM, MINIMUM, COUNT} e  $A_i$  é um atributo de R.

- A relação resultante da operação de agregação tem apenas os atributos definidos em ATRIBS mais um atributo por cada item da lista FUNS, cujo nome resulta da concatenação do nome da função de agregação com o nome do atributo de R.

## Operação de Agregação

- Obtenha o valor do salário máximo, do salário mínimo e da soma do salário de todos os empregados.

$$\mathcal{F} \text{ MAX Salário, MIN Salário, SUM Salário}(\text{EMPREGADO})$$

- Obtenha o número de empregados por departamento e a respectiva média salarial.

$$\text{NumDep } \mathcal{F} \text{ COUNT NumBI, AVERAGE Salário}(\text{EMPREGADO})$$

NumDep	COUNT_NumBI	AVERAGE_Salário
4	2	5500
1	1	2000
2	1	1500

- Obtenha o nome dos empregados com dois ou mais dependentes.

$$\text{DEPS\_COUNT} \leftarrow \rho_{(\text{NumBI, NumDeps})}(\text{EmpBI } \mathcal{F} \text{ COUNT Nome}(\text{DEPENDENTE}))$$

$$\text{RESULT} \leftarrow \pi_{\text{NomeP, NomeF}}(\sigma_{\text{NumDeps} \geq 2}(\text{DEPS\_COUNT}) * \text{EMPREGADO})$$

## Cálculo Relacional

- Modelo formal que se baseia na lógica de predicados e que permite manipular relações no modelo relacional.
- O cálculo relacional tem o mesmo poder expressivo da álgebra relacional. Uma expressão do cálculo relacional é igualmente uma relação que representa o resultado de uma consulta à base de dados.
- As expressões do cálculo podem ser especificadas em termos de variáveis sobre os tuplos, **cálculo relacional por tuplos (ou CRT)**, ou em termos de variáveis sobre o domínio dos atributos, **cálculo relacional por domínios (ou CRD)**.
- O cálculo relacional é uma linguagem **não-procedimental**. Nas expressões do cálculo não se especifica o modo de obter o resultado mas sim o tipo de informação que se pretende obter. Isto difere da álgebra relacional onde é necessário especificar a sequência de operações a aplicar para obter o resultado.
- A linguagem SQL baseia-se em parte no cálculo relacional por tuplos.

## Cálculo Relacional por Tuplos

- O CRT é baseado na especificação de variáveis sobre os tuplos, onde cada variável pode tomar como valor qualquer um dos tuplos de uma determinada relação.
- A forma mais simples de uma expressão do CRT é
 
$$\{t \mid \text{COND}(t)\}$$
 em que  $t$  é uma variável que representa os tuplos de uma relação e  $\text{COND}(t)$  é uma condição sobre  $t$ . O resultado desta expressão é o conjunto dos tuplos  $t$  que satisfazem  $\text{COND}(t)$ .
- A forma genérica de uma expressão do CRT é
 
$$\{t_1.A_1, t_2.A_2, \dots, t_n.A_n \mid \text{COND}(t_1, t_2, \dots, t_n)\}$$
 em que  $t_1, t_2, \dots, t_n$  são variáveis que representam tuplos de relações, cada  $A_i$  é um atributo da relação na qual  $t_i$  toma valores e  $\text{COND}(t_1, t_2, \dots, t_n)$  é uma fórmula bem-formada do cálculo.

## Cálculo Relacional por Tuplos

- Uma fórmula do CRT pode ser formada por condições da forma

$$R(t_i) \text{ ou } t_i.A \theta t_j.B \text{ ou } t_i.A \theta \text{VAL}$$

em que R é uma relação,  $t_i$  e  $t_j$  são variáveis que representam tuplos de relações, A e B são atributos das relações nas quais  $t_i$  e  $t_j$  tomam valores,  $\theta$  é um operador de comparação  $\{=, <, >, \leq, \geq, \neq\}$  e  $\text{VAL} \in \text{dom}(t_i.A)$ .

- Uma fórmula do CRT também pode ser composta por várias fórmulas ligadas pelos operadores lógicos AND, OR ou NOT

$$F_1 \text{ AND } F_2 \text{ ou } F_1 \text{ OR } F_2 \text{ ou } \text{NOT } F_1$$

em que  $F_1$  e  $F_2$  são fórmulas.

## Cálculo Relacional por Tuplos

- Obtenha os empregados que trabalham no departamento 4.

$$\{e \mid \text{EMPREGADO}(e) \text{ AND } e.\text{NumDep} = 4\}$$

Na álgebra relacional seria:

$$\sigma_{\text{NumDep}=4}(\text{EMPREGADO})$$

- Obtenha o número do BI dos empregados que trabalham no departamento 4 e cujo salário é superior a 2000 euros.

$$\{e.\text{NumBI} \mid \text{EMPREGADO}(e) \text{ AND } e.\text{NumDep} = 4 \text{ AND } e.\text{Salário} > 2000\}$$

Na álgebra relacional seria:

$$\pi_{\text{NumBI}}(\sigma_{\text{NumDep}=4 \text{ AND } \text{Salário} > 2000}(\text{EMPREGADO}))$$

## Quantificadores

- Nas fórmulas do CRT podemos ainda usar os quantificadores universal e existencial

$$(\forall t)(F) \quad \text{ou} \quad (\exists t)(F)$$

em que  $\forall$  e  $\exists$  são os quantificadores universal e existencial,  $t$  é uma variável que representa os tuplos de uma relação e  $F$  é uma fórmula.

- Uma variável que representa os tuplos de uma relação diz-se livre se não estiver quantificada. Por exemplo, na fórmula

$$(\exists t)(t.\text{EmpBI} = u.\text{NumBI})$$

$t$  é uma variável quantificada e  $u$  é uma variável livre.

- Numa expressão do CRT, as únicas variáveis que devem aparecer livres são as que aparecem à esquerda da barra |.

## Quantificadores

- $(\forall t)(F)$  é verdadeiro se  $F$  for verdadeiro para todos os tuplos  $t$  em que  $t$  é uma variável livre de  $F$ . Caso contrário  $(\forall t)(F)$  é falso.
- $(\exists t)(F)$  é verdadeiro se  $F$  for verdadeiro para algum tuplo  $t$  em que  $t$  é uma variável livre de  $F$ . Caso contrário  $(\exists t)(F)$  é falso.
- O quantificador universal pode ser transformado no existencial e vice-versa:

$$(\forall t)(F) \equiv \text{NOT } (\exists t)(\text{NOT } F)$$

$$(\exists t)(F) \equiv \text{NOT } (\forall t)(\text{NOT } F)$$

## Quantificadores

- Obtenha o nome dos empregados que trabalham no departamento de Produção.  
 $\{e.NomeP, e.NomeF \mid EMPREGADO(e) \text{ AND } (\exists d)(DEPARTAMENTO(d) \text{ AND } d.Nome = 'Produção' \text{ AND } d.Num = e.NumDep)\}$
- Para todos os projectos localizados no Porto, obtenha o nome do projecto e o último nome do respectivo gerente.  
 $\{p.Nome, g.NomeF \mid PROJECTO(p) \text{ AND } p.Localização = 'Porto' \text{ AND } EMPREGADO(g) \text{ AND } (\exists d)(DEPARTAMENTO(d) \text{ AND } d.Num = p.NumDep \text{ AND } d.GerenteBI = g.NumBI)\}$
- Obtenha o nome dos empregados que trabalham em pelo menos um projecto controlado pelo departamento 4.  
 $\{e.NomeP, e.NomeF \mid EMPREGADO(e) \text{ AND } (\exists p)(\exists t)(PROJECTO(p) \text{ AND } p.NumDep = 4 \text{ AND } TRABALHA\_EM(t) \text{ AND } t.EmpBI = e.NumBI \text{ AND } p.Num = t.NumProj)\}$

## Quantificadores

- Obtenha o número do BI dos empregados que não têm dependentes.  
 $\{e.NumBI \mid EMPREGADO(e) \text{ AND } \text{NOT } (\exists d)(DEPENDENTE(d) \text{ AND } d.EmpBI = e.NumBI)\}$   
 Alternativa de resolução utilizando a regra de transformação:  
 $\{e.NumBI \mid EMPREGADO(e) \text{ AND } (\forall d)(\text{NOT } DEPENDENTE(d) \text{ OR } d.EmpBI \neq e.NumBI)\}$
- Obtenha o nome dos empregados que trabalham em todos os projectos controlados pelo departamento 4.  
 $\{e.NomeP, e.NomeF \mid EMPREGADO(e) \text{ AND } (\forall p)(\text{NOT } PROJECTO(p) \text{ OR } p.NumDep \neq 4 \text{ OR } (\exists t)(TRABALHA\_EM(t) \text{ AND } t.EmpBI = e.NumBI \text{ AND } t.NumProj = p.Num))\}$

### Cálculo Relacional por Domínios

- O CRD é baseado na especificação de variáveis sobre o domínio dos atributos, onde cada variável pode tomar como valor os valores de um determinado domínio.

- A forma genérica de uma expressão do CRD é

$$\{v_1, v_2, \dots, v_n \mid \text{COND}(v_1, v_2, \dots, v_n, v_{n+1}, v_{n+2}, \dots, v_{n+m})\}$$

em que  $v_1, v_2, \dots, v_n, v_{n+1}, v_{n+2}, \dots, v_{n+m}$  são variáveis que representam valores de domínios e  $\text{COND}(v_1, v_2, \dots, v_n, v_{n+1}, v_{n+2}, \dots, v_{n+m})$  é uma fórmula bem-formada do cálculo.

- O resultado desta expressão é a relação de grau  $n$  que corresponde às variáveis  $v_1, v_2, \dots, v_n$  e cujos valores satisfazem  $\text{COND}(v_1, v_2, \dots, v_n, v_{n+1}, v_{n+2}, \dots, v_{n+m})$ .

### Cálculo Relacional por Domínios

- Uma fórmula do CRD pode ser formada por condições da forma

$$R(v_1, v_2, \dots, v_n) \quad \text{ou} \quad v_i \theta v_j \quad \text{ou} \quad v_i \theta \text{VAL}$$

em que  $R$  é uma relação, cada  $v_i$  é uma variável que representa valores do domínio do atributo  $A_i$  de  $R$ ,  $\theta$  é um operador de comparação  $\{=, <, >, \leq, \geq, \neq\}$  e  $\text{VAL} \in \text{dom}(A_i)$ .

- Uma fórmula do CRD também pode ser composta por várias fórmulas ligadas pelos operadores lógicos AND, OR ou NOT ou usar os quantificadores universal e existencial

$$F_1 \text{ AND } F_2 \quad \text{ou} \quad F_1 \text{ OR } F_2 \quad \text{ou} \quad \text{NOT } F_1$$

$$(\forall v)(F_1) \quad \text{ou} \quad (\exists v)(F_1)$$

em que  $F_1$  e  $F_2$  são fórmulas,  $\forall$  e  $\exists$  são os quantificadores universal e existencial e  $v$  é uma variável que representa valores de um domínio.

### Cálculo Relacional por Domínios

- Obtenha a data de aniversário do empregado Rui Silva.

$$\{g \mid (\exists a)(\exists b)(\exists c)(\exists d)(\exists e)(\exists f)(\exists h)(\exists i)(\text{EMPREGADO}(a,b,c,d,e,f,g,h,i) \text{ AND } a = \text{'Rui'} \text{ AND } b = \text{'Silva'})\}$$

Simplificação de notação: não utilizar vírgulas para separar as variáveis nas relações e assumir que as variáveis que não aparecem em qualquer condição são quantificadas existencialmente.

$$\{g \mid (\exists a)(\exists b)(\text{EMPREGADO}(abcdefghi) \text{ AND } a = \text{'Rui'} \text{ AND } b = \text{'Silva'})\}$$

Simplificação de notação: utilizar valores directamente nas relações e assumir que as variáveis que não aparecem à esquerda da barra são quantificadas existencialmente.

$$\{g \mid \text{EMPREGADO}(\text{'Rui'}, \text{'Silva'}, c,d,e,f,g,h,i)\}$$

### Cálculo Relacional por Domínios

- Para todos os projectos localizados no Porto, obtenha o nome do projecto e o último nome do respectivo gerente.

$$\{pb \mid (\exists m)(\exists n)(\exists r)(\exists s)(\text{EMPREGADO}(abcdefghi) \text{ AND DEPARTAMENTO}(lmno) \text{ AND PROJECTO}(pqrs) \text{ AND } r = \text{'Porto'} \text{ AND } s = m \text{ AND } n = c)\}$$

- Obtenha o número do BI dos empregados que não têm dependentes.

$$\{c \mid \text{EMPREGADO}(abcdefghi) \text{ AND NOT } (\exists v)(\text{DEPENDENTE}(vwxyz) \text{ AND } c = v)\}$$

Alternativa de resolução utilizando a simplificação de notação:

$$\{c \mid \text{EMPREGADO}(abcdefghi) \text{ AND NOT DEPENDENTE}(cwxyz)\}$$

## Álgebra e Cálculo Relacional

- Considere as relações  $R(A,B,C,D)$  e  $S(A,B)$ . Escreva uma expressão no cálculo relacional por tuplos e no cálculo relacional por domínios para a seguinte expressão da álgebra relacional:

$$\pi_{A,B}(R) \cap S$$

- CRT

$$\{r.A, r.B \mid R(r) \text{ AND } (\exists s)(S(s) \text{ AND } r.A = s.A \text{ AND } r.B = s.B)\}$$

- CRD

$$\{ab \mid R(abcd) \text{ AND } (\exists e)(\exists f)(S(ef) \text{ AND } a = e \text{ AND } b = f)\}$$

$$\{ab \mid R(abcd) \text{ AND } S(ab)\}$$

## Álgebra e Cálculo Relacional

- Considere as relações  $R(A,B,C,D)$  e  $S(A,B)$ . Escreva uma expressão no cálculo relacional por tuplos e no cálculo relacional por domínios para a seguinte expressão da álgebra relacional:

$$R * (\pi_{A,B}(R) - S)$$

- CRT

$$\{r \mid R(r) \text{ AND NOT } (\exists s)(S(s) \text{ AND } r.A = s.A \text{ AND } r.B = s.B)\}$$

- CRD

$$\{abcd \mid R(abcd) \text{ AND NOT } (\exists e)(\exists f)(S(ef) \text{ AND } a = e \text{ AND } b = f)\}$$

$$\{abcd \mid R(abcd) \text{ AND NOT } S(ab)\}$$

## Álgebra e Cálculo Relacional

- Considere as relações R(A,B,C,D) e S(A,B). Escreva uma expressão no cálculo relacional por tuplos e no cálculo relacional por domínios para a seguinte expressão da álgebra relacional:

$$\sigma_{C=D}((R * \pi_A(S)) * \pi_B(S))$$

- **CRT**

$$\{r \mid R(r) \text{ AND } (\exists s_1)(S(s_1) \text{ AND } r.A = s_1.A) \\ \text{AND } (\exists s_2)(S(s_2) \text{ AND } r.B = s_2.B) \text{ AND } r.C = r.D\}$$

- **CRD**

$$\{abcd \mid R(abcd) \text{ AND } (\exists e)(S(ef) \text{ AND } e = a) \\ \text{AND } (\exists h)(S(gh) \text{ AND } h = b) \text{ AND } c = d\} \\ \{abcc \mid R(abcc) \text{ AND } S(ax) \text{ AND } S(yb)\}$$

## Álgebra e Cálculo Relacional

- Considere as relações R(A,B,C,D) e S(A,B). Escreva uma expressão no cálculo relacional por tuplos e no cálculo relacional por domínios para a seguinte expressão da álgebra relacional:

$$\pi_{A,B,D}(R) \div S$$

- **CRT**

$$\{r_1.D \mid R(r_1) \text{ AND } (\forall s)(\text{NOT } S(s) \text{ OR } (\exists r_2) \\ (R(r_2) \text{ AND } s.A = r_2.A \text{ AND } s.B = r_2.B \text{ AND } r_1.D = r_2.D))\} \\ \{r_1.D \mid R(r_1) \text{ AND NOT } (\exists s)(S(s) \text{ AND } (\forall r_2) \\ (\text{NOT } R(r_2) \text{ OR } s.A \neq r_2.A \text{ OR } s.B \neq r_2.B \text{ OR } r_1.D \neq r_2.D))\}$$

- **CRD**

$$\{d \mid R(abcd) \text{ AND } (\forall e)(\forall f)(\text{NOT } S(ef) \text{ OR } R(efxd))\}$$