

## Álgebra Relacional

A álgebra relacional é uma linguagem procedural, que possui uma coleção de operações que são utilizadas para manipular relações (tabelas) inteiras. O resultado de cada operação é uma nova relação, que pode ser manipulada posteriormente.

As operações de álgebra relacional são, usualmente, divididas em dois grupos:

⇒ Operações da teoria de conjuntos da matemática: UNIÃO, INTERSEÇÃO, DIFERENÇA e PRODUTO.

⇒ O outro grupo consiste de operações desenvolvidas especificamente para banco de dados relacionais, que inclui SELEÇÃO, PROJEÇÃO, JUNÇÃO, entre outras.

A tabela a seguir será usada para exemplificar as operações:

Nome	Sobren	RG	DataNas	Endereço	Sex	Salario	RGSuper	NroDepto
João	Silva	123456789	09/01/55	Rua 15 de Novembro, 731, São Carlos	M	3.000	333445555	5
Francisco	Souza	333445555	08/12/45	Rua Jorge Assef, 100, São Paulo	M	4.000	888665555	5
Alice	Fernandes	999887777	18/03/75	Rua 9 de Julho, 535, São Carlos	F	2.500	987654321	4
Joana	Pereira	987654321	20/06/31	Rua Tiradentes, 291, Ri. Preto	F	4.300	888665555	4
Rodolfo	Nogueira	666884444	15/09/52	Av. Brasil, 1005, Fernandópolis	M	3.800	333445555	5

### 1. A OPERAÇÃO SELEÇÃO(SELECT)

É utilizada para selecionar um subconjunto de tuplas (registros) numa relação (tabela) que satisfaça uma **condição de seleção**. Por exemplo, para selecionar o subconjunto de registros de EMPREGADO que trabalha no departamento 4 ou cujo salário é maior do que R\$ 3.000,00, pode-se especificar cada uma dessas duas condições com a operação SELECT, como segue:

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

$$\sigma_{\text{Salario} > 3000} (\text{EMPREGADO})$$

Em geral, a operação SELECT é denotada por:

$$\sigma_{\langle \text{condição da seleção} \rangle} (\text{Nome da Relação})$$

onde o símbolo  $\sigma$  (sigma) é usado para denotar o operador SELECT e a condição de seleção é uma expressão booleana especificada nos atributos da relação.

A relação resultante da operação SELECT tem os mesmos atributos da relação especificada em <Nome da Relação>. A expressão booleana especificada em <condição da seleção> tem a forma:

<Nome do Atributo> <operador de comparação> <valor constante>, ou  
 <Nome do Atributo> <operador de comparação> <Nome do Atributo>

onde <Nome do Atributo> é o nome de um atributo de <Nome da Relação>, <operador de comparação> é normalmente um dos seguintes operadores {=, <, >, ≥, ≤, ≠} e <valor constante> é um valor constante do domínio do atributo. As cláusulas podem ser arbitrariamente conectadas pelos operadores booleanos, AND, OR e NOT, para formar uma condição de seleção geral.

Por exemplo, para selecionar as tuplas de todos os empregados que trabalham no departamento 4 e recebem mais de R\$ 4.000,00, ou trabalham no departamento 5 e recebem mais de R\$ 3.000,00, podemos especificar a seguinte operação.

$\sigma$  (NroDepto = 4 AND Salario > 4000) OR (NroDepto = 5 AND Salario > 3000) (EMPREGADO)

O resultado está exibido na tabela a seguir.

Nome	Sobren	RG	DataNas	Endereço	Sex	Salario	RGSuper	NroDepto
Francisco	Souza	333445555	08/12/45	Rua Jorge Assef, 100, São Paulo	M	4.000	888665555	5
Joana	Pereira	987654321	20/06/31	Rua Tiradentes, 291, Ri. Preto	F	4.300	888665555	4
Rodolfo	Nogueira	666884444	15/09/52	Av. Brasil, 1005, Fernandópolis	M	3.800	333445555	5

Note que os operadores de comparação {=, <, >, ≥, ≤, ≠} aplicam-se a atributos cujos domínios são *valores ordenados*, tais como domínios de números ou datas. Se o domínio de um atributo é um conjunto de *valores desordenados*, então somente os operadores de comparação {=, ≠} podem ser aplicados ao atributo. Um exemplo de um domínio desordenado é o domínio Cor = {vermelho, azul, verde, branco, amarelo, ...}.

O grau de uma relação resultante de uma operação SELECT é o mesmo da relação R original na qual a operação é aplicada, pois possui os mesmos atributos de R. O número de tuplas na relação resultante é sempre menor ou igual ao número de tuplas da relação original R.

A operação SELECT é **comutativa**, isto é,  $\sigma_{<cond1>}(\sigma_{<cond2>}(R)) = \sigma_{<cond2>}(\sigma_{<cond1>}(R))$ . Assim, uma seqüência de SELECT's pode ser aplicada em qualquer ordem. Além disso, podemos sempre combinar uma **cascata** de operações SELECT em uma única operação SELECT com uma condição conjuntiva (AND), isto é:

$$\sigma_{<cond1>}(\sigma_{<cond2>}(\dots((\sigma_{<condn>}(R))\dots))) = \sigma_{<cond1>} \wedge <cond2> \wedge <condn> (R)$$

## 2. A OPERAÇÃO PROJEÇÃO(PROJECT)

A operação PROJECT seleciona certas *colunas* da tabela e descarta as outras colunas. Exemplo: listar o nome, sobrenome e o salário de cada empregado. Pode-se usar a operação PROJECT da seguinte forma:

$$\pi_{\text{Nome, Sobren, Salario}}(\text{EMPREGADO})$$

A tabela a seguir mostra o resultado da operação  $\pi_{\text{Nome, Sobren, Salario}}(\text{EMPREGADO})$ :

Nome	Sobren	Salario
João	Silva	3.000
Francisco	Souza	4.000
Alice	Fernandes	2.500
Joana	Pereira	4.300
Rodolfo	Nogueira	3.800

A forma geral da operação PROJECT é:

$$\pi_{\langle \text{lista de atributos} \rangle}(\text{Nome da Relação})$$

onde  $\pi$  (pi) é o símbolo usado para representar a operação PROJECT e  $\langle \text{lista de atributos} \rangle$  é uma lista de atributos da relação especificada por  $\langle \text{Nome da Relação} \rangle$ . A relação resultante tem somente os atributos especificados em  $\langle \text{lista de atributos} \rangle$  e *na mesma ordem que aparecem na lista*. Assim, seu **grau** é igual ao número de atributos em  $\langle \text{lista de atributos} \rangle$ .

O número de tuplas em uma relação resultante de uma operação PROJECT é sempre menor ou igual ao número de tuplas da relação original.

A propriedade **comutativa** não é mantida em PROJECT.

## 3. SEQÜÊNCIAS DE OPERAÇÕES E RENOMEAÇÃO DE ATRIBUTOS

As operações podem ser aplicadas de forma aninhada, usando uma única expressão da álgebra relacional, como a seguir, cuja expressão recupera o primeiro nome, o último nome e o salário de todos os empregados que trabalham no departamento cujo código é 5:

$$\pi_{\text{Nome, Sobren, Salario}}(\sigma_{\text{NroDepto} = 5}(\text{EMPREGADO}))$$

A tabela a seguir mostra o resultado da operação anterior:

Nome	Sobren	Salario
João	Silva	3.000
Francisco	Souza	4.000
Rodolfo	Nogueira	3.800

Podemos ainda criar relações resultantes intermediárias, dando um nome para cada relação intermediária:

$$R1 \leftarrow \sigma_{\text{NroDepto} = 5}(\text{EMPREGADO})$$

$$\text{RESULT} \leftarrow \pi_{\text{Nome, Sobren, Salario}}(R1)$$

Podemos também utilizar esta técnica para renomear os atributos nas relações intermediárias e resultantes, como a seguir. Isso pode ser útil na conexão com operações mais complexas, tais como UNION e JOIN.

$$TEMP \leftarrow \sigma_{NroDepto = 5}(EMPREGADO)$$

$$R(FirstName, LastName, Salary) \leftarrow \pi_{Nome, Sobren, Salario}(TEMP)$$

A tabela a seguir mostra o resultado da operação  $TEMP \leftarrow \sigma_{NroDepto = 5}(EMPREGADO)$

**TEMP**

Nome	Sobren	RG	DataNas	Endereço	Sex	Salario	RGSuper	NroDepto
João	Silva	123456789	09/01/55	Rua 15 de Novembro, 731, São Carlos	M	3.000	333445555	5
Francisco	Souza	333445555	08/12/45	Rua Jorge Assef, 100, São Paulo	M	4.000	888665555	5
Rodolfo	Nogueira	666884444	15/09/52	Av. Brasil, 1005, Fernandópolis	M	3.800	333445555	5

A tabela a seguir mostra o resultado da operação  $R(FirstName, LastName, Salary) \leftarrow \pi_{Nome, Sobren, Salario}(TEMP)$

**R**

FirstName	LastName	Salary
João	Silva	3.000
Francisco	Souza	4.000
Rodolfo	Nogueira	3.800

#### 4. OPERAÇÕES DA TEORIA DOS CONJUNTOS

O próximo grupo de operações da álgebra relacional são operações matemáticas, padrões em conjuntos. Elas se aplicam ao modelo relacional, pois uma relação é definida como um conjunto de tuplas e pode ser usada para processar as tuplas em duas relações como conjuntos. Por exemplo, para recuperar os RG's de todos os empregados que trabalham no departamento 5 ou supervisione diretamente um empregado que trabalhe no departamento 5, podemos utilizar a operação UNION:

$$DEP_5EMPS \leftarrow \sigma_{NroDepto = 5}(EMPREGADO)$$

$$RESULT \leftarrow \pi_{RG}(DEP_5EMPS)$$

$$RESULT1(rg) \leftarrow \pi_{RGSuper}(DEP_5EMPS)$$

$$RESULT2 \leftarrow RESULT1 \cup RESULT2$$

A relação RESULT1 tem os RG's de todos os empregados que trabalham no departamento 5, enquanto que RESULT2 tem os RG's de todos os empregados que supervisionam diretamente um empregado que trabalha no departamento 5. A operação UNION produz as tuplas que ou estão em RESULT1 ou estão em RESULT2 ou estão em ambas. O resultado da operação  $RESULT \leftarrow RESULT1 \cup RESULT2$  está apresentado a seguir:

<b>RESULT</b>	<table border="1"><tr><th>RG</th></tr><tr><td>123456789</td></tr><tr><td>333445555</td></tr><tr><td>666884444</td></tr></table>	RG	123456789	333445555	666884444	<b>RESULT1</b>	<table border="1"><tr><th>RG</th></tr><tr><td>333445555</td></tr><tr><td>888665555</td></tr><tr><td>333445555</td></tr></table>	RG	333445555	888665555	333445555	<b>RESULT2</b>	<table border="1"><tr><th>RG</th></tr><tr><td>123456789</td></tr><tr><td>333445555</td></tr><tr><td>666884444</td></tr><tr><td>888665555</td></tr></table>	RG	123456789	333445555	666884444	888665555
RG																		
123456789																		
333445555																		
666884444																		
RG																		
333445555																		
888665555																		
333445555																		
RG																		
123456789																		
333445555																		
666884444																		
888665555																		

Para a utilização dessas operações em banco de dados relacionais devemos garantir que o resultado da aplicação dessas operações em duas relações gere uma terceira relação também válida. Para isso, as duas relações envolvidas devem possuir o **mesmo tipo de tuplas**. Essa condição é chamada de *compatibilidade de união*.

Duas relações  $R(A_1, A_2, A_3, \dots, A_n)$  e  $S(B_1, B_2, B_3, \dots, B_n)$  são ditas **união-compatíveis** se possuem o **mesmo grau n**, e se  $\text{dom}(A_i) = \text{dom}(B_i)$  para  $1 \leq i \leq n$ . Isto significa que as duas relações possuem o mesmo número de atributos e que cada par de atributos correspondentes possuem o mesmo domínio. Obs: alguns autores consideram ser necessário também que, as colunas das duas relações tenham os mesmos nomes.

As duas relações a seguir são união-compatíveis.

#### Estudante

N	S
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

#### Instrutor

Nome	Sobrenome
John	Smith
Ricardo	Browne
Susan	Yao
Francis	Johnson
Ramesh	Shah

Podemos definir as operações UNION, INTERSECTION e DIFFERENCE em duas relações união-compatíveis R e S, como se segue:

- 4.1) UNION:** O resultado desta operação, denotada por  $R \cup S$ , é uma relação que inclui todas as tuplas que estão em R ou em S ou em ambas R e S. Tuplas duplicadas são eliminadas. A tabela a seguir mostra o resultado da operação ESTUDANTE  $\cup$  INSTRUTOR.

N	S
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert
John	Smith
Ricardo	Browne
Francis	Johnson

#### União Exclusiva

A união exclusiva de duas relações R e S, resulta em uma relação que contém as tuplas que estão em S ou em R, mas não as tuplas que estão em ambas. A representação utilizada para a operação União Exclusiva é:

$$R \cup | S$$

**4.2) INTERSECTION:** O resultado desta operação, denotada por  $R \cap S$ , é uma relação que inclui todas as tuplas que estão em ambas R e S. A tabela a seguir mostra o resultado da operação ESTUDANTE  $\cap$  INSTRUTOR.

N	S
Susan	Yao
Ramesh	Shah

Obs: Ambas as operações UNION e INTERSECTION:

- ⇒ são comutativas, ou seja,  $R \cup S = S \cup R$  e  $R \cap S = S \cap R$
- ⇒ são aplicadas a qualquer número de relações
- ⇒ são associativas, ou seja,  $R \cup (S \cup T) = (R \cup S) \cup T$  e  $R \cap (S \cap T) = (R \cap S) \cap T$

**4.3) DIFFERENCE:** O resultado dessa operação, denotada por  $R - S$ , é uma relação que inclui todas as tuplas que estão em R, mas não estão em S. A tabela (a) mostra o resultado da operação ESTUDANTE - INSTRUTOR e a tabela (b) mostra o resultado da operação INSTRUTOR - ESTUDANTE

**(a)**

N	S
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

**(b)**

Nome	Sobrenome
John	Smith
Ricardo	Browne
Francis	Johnson

A operação DIFFERENCE não é comutativa pois em geral,  $R - S \neq S - R$ . Note que a convenção adotada é que a relação resultante tem os mesmos nomes de atributos da primeira relação R.

**4.4) PRODUTO CARTESIANO (X)**

Combina tuplas de duas relações R e S, resultando em uma relação que tem uma tupla para cada combinação de tuplas - uma de R e outra de S:

$$Q(A_1, A_2, A_3, \dots, A_n, B_1, B_2, B_3, \dots, B_m) \leftarrow R(A_1, A_2, A_3, \dots, A_n) \times S(B_1, B_2, B_3, \dots, B_m)$$

Exemplo:

**R1**

Nome	Sobren
João	Silva
Francisco	Souza

**R2**

NroDepto	NomeDepto
1	Administração
2	Finanças

**R1 X R2**

Nome	Sobren	NroDepto	NomeDepto
João	Silva	1	Administração
João	Silva	2	Finanças
Francisco	Souza	1	Administração
Francisco	Souza	2	Finanças

#### 4.5) JOIN - Operação Junção (X)

Denotada por  $|X|$ , é utilizada para combinar tuplas de duas relações através de um ou mais atributos comuns às duas relações. A tabela resultante contém as colunas das duas tabelas que participaram da junção:

$$Q (A_1, A_2, A_3, \dots, A_n, B_1, B_2, B_3, \dots, B_m) \leftarrow R (A_1, A_2, A_3, \dots, A_n) |X| S(B_1, B_2, B_3, \dots, B_m)$$

Esta operação é muito importante para qualquer banco de dados relacional com mais do que uma relação, pois nos permite processar relacionamento entre relações.

A forma geral de uma operação JOIN em 2 relações  $R (A_1, A_2, \dots, A_n)$  e  $S (B_1, B_2, \dots, B_m)$  é:

$$R |X| \langle \text{condição de junção} \rangle S$$

Uma condição de junção é da forma:  $\langle \text{condição} \rangle \text{ AND } \langle \text{condição} \rangle \text{ AND } \dots \text{ AND } \langle \text{condição} \rangle$ , onde a condição é da forma  $A_i \theta B_j$ , A e B são atributos com o mesmo domínio e  $\theta$  é um dos operadores de comparação  $\{=, <, \leq, >, \geq, \neq\}$ .

A operação de junção mais comumente usada envolve, na condição JOIN, somente comparações de igualdade. Nesse caso ela é chamada EQUIJOIN. Essa operação conserva pares de atributos com valores idênticos em cada tupla. A operação **Natural Join (notação \*)** conserva apenas uma das colunas de valores repetidos, mas para que seja possível, é necessário que os atributos da condição tenham o mesmo nome em ambas as relações.

No exemplo a seguir é utilizada a EQUIJOIN:

EMPREGADO				DEPARTAMENTO		
Nome	Sobren	RG	NroDepto	NroDepto	NomeDepto	RGGerente
João	Silva	123456789	1	1	Pesquisa	333445555
Francisco	Souza	333445555	1	2	Administração	987654321
Alice	Fernandes	999887777	4	3	Finanças	666884444
Joana	Pereira	987654321	2	4	Pessoal	999887777
Rodolfo	Nogueira	666884444	3			

Suponha que se queira recuperar o nome do gerente de cada departamento. Para obter o nome do gerente, precisamos combinar cada tupla de Departamento com a tupla de Empregado cujo valor do RG seja igual ao valor de RGGerente na tupla de Departamento.

$$\text{DEPT\_GER} \leftarrow \text{DEPARTAMENTO} |X|_{\text{RGGerente} = \text{RG}} (\text{EMPREGADO})$$

$$\text{RESULT} \leftarrow \pi_{\text{NroDepto, Nome, Sobren}} (\text{DEPT\_GER})$$

DEPT_GER						
NroDepto	NomeDepto	RGGerente	Nome	Sobren	RG	NroDepto
1	Pesquisa	333445555	Francisco	Souza	333445555	1
2	Administração	987654321	Joana	Pereira	987654321	2
3	Finanças	666884444	Rodolfo	Nogueira	666884444	3
4	Pessoal	999887777	Alice	Fernandes	999887777	4

#### RESULT

NroDepto	Nome	Sobren
1	Francisco	Souza
2	Joana	Pereira
3	Rodolfo	Nogueira
4	Alice	Fernandes

**Exercício:** Formule a seguinte consulta na álgebra relacional: Recupere o nome e o endereço de todos os empregados que trabalham para o departamento "Pesquisa"

#### 4.6) OPERAÇÃO DIVISÃO (DIVISION)

A operação DIVISION é útil para um tipo especial de consulta que algumas vezes ocorrem em aplicações de banco de dados. Esta operação aplicada sobre duas relações R e S, produz uma relação Q que inclui todas as tuplas que aparecem em R combinadas com todas as tuplas de S. A representação utilizada por esta operação é:  $R \div S$

Como exemplo, tem-se: "Recupere os números dos empregados que trabalham em *todos* os projetos".

##### PROJETO

NomeProj	NroProj	NroDepto
PROJETO A	P1	1
PROJETO B	P2	3
PROJETO c	P3	1

##### TRABALHA

NroEmp	NroProj	Horas
1	P1	10
1	P3	15
2	P1	10
2	P2	20
2	P3	15
3	P2	30
4	P1	10
4	P2	10
4	P3	20

Para expressar essa consulta utilizando DIVISION, realizamos as seguintes operações:

$$R \leftarrow \pi_{\text{NroProj}}(\text{PROJETO})$$

$$S \leftarrow \pi_{\text{NroEmp, NroProj}}(\text{TRABALHA})$$

$$\text{RESULT} \leftarrow R \div S$$

NroProj
P1
P2
P3

NroEmp	NroProj
1	P1
1	P3
2	P1
2	P2
2	P3
3	P2
4	P1
4	P2
4	P3

NroEmp
2
4



**Exercícios:**

1- Faça a expressão algébrica na Álgebra Relacional das seguintes consultas:

OBS: Instancie a relação DEPENDENTE cujo esquema da relação é:  
DEPENDENTE (RG, Nome\_Depend, Sexo, DtaNasc, Grau\_Parent)

- A - Consulte os nomes dos empregados que trabalham em todos os projetos controlados pelo departamento 5.
- B - Faça uma lista de número de projetos para projetos que envolvem um empregado cujo sobrenome é 'Silva' como trabalhador ou como gerente do departamento que controla o projeto.
- C - Liste o nome de todos os empregados que não possuam dependentes.
- D - Liste os nomes dos gerentes que possuem no mínimo um dependente.

2- Dado o seguinte esquema de Banco de Dados:  $\varphi$

*Empregado* (nome\_empregado, rua, cidade);  
*Trabalha* (#nome\_empregado, #nome\_companhia, salário);  
*Companhia* (nome\_companhia, cidade);  
*Ger\_Emp* (#nome\_empregado, #nome\_gerente)

Resolva as consultas abaixo utilizando álgebra Relacional:

- A- Encontre os nomes de todos os empregados que trabalham para a companhia 'UNIP';
- B- Liste todos os nomes das cidades dos empregados que trabalham na companhia 'UNIP';
- C- Recupere o nome, cidade e rua da residência de todos os empregados da companhia UNIP que ganham mais de vinte mil reais por ano – 13° e 1/3 de férias;
- D- Encontre os nomes de todos os empregados que moram na mesma cidade da companhia que trabalham;
- E- Encontre os nomes de todos os empregados que moram na mesma rua e cidade de seu gerente;
- F- Liste os nomes de todos os empregados que não trabalham para a companhia 'SBD';
- G- Assuma que as companhias possam estar localizadas em diversas cidades. Encontre todas as companhias localizadas em todas as cidades onde haja unidades da companhia 'SBD'.