

Álgebra Relacional

Álgebra Relacional é o conjunto básico de operações para o Modelo Relacional. Essas operações permitem a um usuário especificar as solicitações básicas de recuperação.

O resultado de uma recuperação será uma nova relação que pode ter sido formada de uma ou mais relações.

Uma sequência de operações de Álgebra Relacional forma uma expressão de Álgebra Relacional cujos resultados também serão uma relação que representa o resultado de uma consulta de BD (ou solicitação de recuperação).

Importância da Álgebra Relacional:

- Provê um fundamento formal para operações do modelo relacional;
- É usada como uma base para implementar e otimizar as consultas em SGBDRs;
- Alguns de seus conceitos são incorporados na linguagem de consulta padrão SQL para SGBDRs.

Operadores definidos pelo modelo relacional:

Seleção	União
Projeção	Interseção
Junção	Diferença
Divisão	Produto Cartesiano

OBS: as operações Seleção e Projeção são operações unárias porque atuam em relações únicas.

PROFESSOR

Matricula	Nome
101	Ana
110	Carlos
120	Jaqueline

ALUNO

Matricula	Nome
103	Marcos
101	Ana
105	Janete

União

$$A \cup B$$

União é a operação entre duas relações cujo resultado contém todas as tuplas presentes em uma das relações originais ou em ambas, onde:

- As relações devem ter uma estrutura similar, com o mesmo número de colunas e domínios compatíveis dos atributos correspondentes.
- Tuplas duplicadas são eliminadas

PROFESSOR \cup ALUNO

Matricula	Nome
101	Ana
110	Carlos
120	Jaqueline
103	Marcos
105	Janete

Interseção

$$A \cap B$$

Interseção é a operação entre duas relações de estrutura similar cujo resultado contém todas as tuplas presentes em ambas.

- Tuplas duplicadas são eliminadas

PROFESSOR \cap ALUNO

Matricula	Nome
101	Ana

Diferença

$A - B$

Diferença é a operação entre duas relações de estrutura similar cujo resultado contém todas as tuplas que estão na primeira relação e não aparecem na segunda.

- $A - B$ é diferente de $B - A$

PROFESSOR - ALUNO

Matricula	Nome
110	Carlos
120	Jaqueline

ALUNO - PROFESSOR

Matricula	Nome
103	Marcos
105	Janete

Seleção

$\sigma_{\text{condição}}(\text{tabela})$

A operação de Seleção é usada para selecionar um subconjunto de tuplas de uma relação que satisfaça uma condição de seleção. Em geral a operação de seleção é indicada por $\sigma_{\langle \text{condição de seleção} \rangle}(R)$. A operação de seleção opera sobre uma relação, gerando uma relação resultante que contém todas as colunas das tuplas selecionadas da relação original.

←

←

←

CIDADE

Cod_Cidade	Nome_Cidade	Capital	UF
1234	Porto Alegre	S	RS
2345	São Paulo	S	SP
1235	Caxias do Sul	N	RS

$\sigma_{UF='RS'}(CIDADE)$

Cod_Cidade	Nome_Cidade	Capital	UF
1234	Porto Alegre	S	RS
1235	Caxias do Sul	N	RS

$\sigma_{UF='RS' \text{ and } Capital='S'}(CIDADE)$

Cod_Cidade	Nome_Cidade	Capital	UF
1234	Porto Alegre	S	RS

Projeção

$\pi_{A_1, A_2, \dots, A_n}(\text{tabela})$

Enquanto a operação de seleção seleciona algumas das tuplas da relação e descarta outras, a operação de projeção seleciona certas colunas da relação e descarta outras.

Se estivermos interessados apenas em certos atributos de uma relação, usamos a operação Projeção.

↓

↓

$\pi_{\text{Nome_Cidade, UF}}(\text{CIDADE})$

Nome_Cidade	UF
Porto Alegre	RS
São Paulo	SP
Caxias do Sul	RS

 $\pi_{\text{UF}}(\text{CIDADE})$

UF
RS
SP
RS

obs: tuplas (linhas) duplicadas são eliminadas

Sequências de Operações e Rename

Normalmente é desejável aplicar várias operações de álgebra relacional, uma após a outra.

Ex.: Listar o primeiro e o último nome dos empregados do departamento 5.

- Temos uma seleção seguida de uma projeção.

1ª opção: Aninhar as operações:

$\pi_{\text{PNAME, UNOME, SALARIO}}(\sigma_{\text{DNO}=5}(\text{EMPREGADO}))$

- Note que a ordem destas operações neste exemplo não pode ser invertida.
- Qual seria a relação de resposta para esta operação? Veja a seguir:

PNAME	UNOME	SALARIO
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000
Joyce	English	25000

2ª opção: criar relações intermediárias:

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

$\text{RESULTADO} \leftarrow \pi_{\text{PNAME, UNOME, SALARIO}}(\text{TEMP})$

- Renomeando atributos durante uma atribuição:

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

$\text{RESULTADO}(\text{PrimeiroNome, UltimoNome, Salario}) \leftarrow \pi_{\text{PNAME, UNOME, SALARIO}}(\text{TEMP})$

O resultado final:

R	PRIMEIRONOME	ULTIMONOME	SALARIO
	John	Smith	30000
	Franklin	Wong	40000
	Ramesh	Narayan	38000
	Joyce	English	25000

Usando o operador **Rename** para renomear atributos e relações.

Suponha que temos uma relação $R(A_1, A_2, \dots, A_n)$.

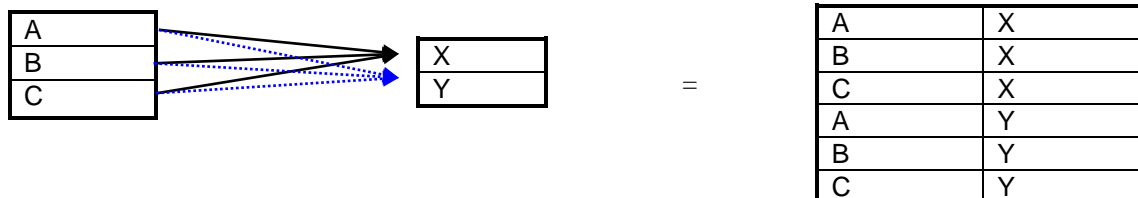
- $\rho(B_1, B_2, \dots, B_n)(R)$: renomeia os atributos de R para B_1, B_2, \dots, B_n .
- $\rho S(R)$: renomeia a relação R (ela passa a se chamar S).
- $\rho S(B_1, B_2, \dots, B_n)(R)$: renomeia a relação em si e os atributos.

Produto Cartesiano

tabela_1 \times tabela_2

A operação Produto Cartesiano é também conhecida como *Cross Product* (Produto Cruzado) ou *Cross Join* (Junção Cruzada) que é indicada por " \times ". Essa é uma operação binária que combina as tuplas de duas relações de forma combinatória.

O resultado de $R(A_1, A_2, \dots, A_n) \times S(B_1, B_2, \dots, B_m)$ é uma relação Q , com grau $n + m$ de atributos $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$, nessa ordem.



Normalmente essa operação aplicada sozinha não tem sentido. Ela será útil quando for seguida por uma seleção que casar os valores de determinados atributos provindos das relações componentes.

Exercícios

1 - Descreva a relação resultado do produto cartesiano $R \times S$

R		S			R \times S				
A	B	B	C	D	A	B	B	C	D
1	2	2	5	6					
3	4	4	7	8					
		9	10	11					

2 - Descreva a relação resultado do produto cartesiano CIDADE \times ESTADO

CIDADE				ESTADO		
Cod_Cidade	Nome_Cid.	Capital	UF	UF	Nome_Estado	Regiao
1234	Porto Alegre	S	RS	RS	Rio Grande do Sul	Sul
2345	São Paulo	S	SP	SP	São Paulo	Sudeste
1235	Caxias do Sul	N	RS			

CIDADE \times ESTADO

3 - Considerando os dados apresentados abaixo, faça:

EMPREGADO	PNOME	MINICIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
	John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
	Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
	Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
	Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
	Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
	Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
	Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
	James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	null	1

DEPT	LOCALIZACOES	DNUMERO	DLOCALIZACAO
		1	Houston
		4	Stafford
		5	Bellaire
		5	Sugarland
			Houston

DEPARTAMENTO	DNO	DNUMERO	GERSSN	GERDATAINICIO
Pesquisa		5	333445555	1988-05-22
Administração		4	987654321	1995-01-01
Sede administrativa		1	888665555	1981-06-19

TRABALHA_EM	ESSN	PNO	HORAS
	123456789	1	32.5
	123456789	2	7.5
	666884444	3	40.0
	453453453	1	20.0
	453453453	2	20.0
	333445555	2	10.0
	333445555	3	10.0
	333445555	10	10.0
	333445555	20	10.0
	999887777	30	30.0
	999887777	10	10.0
	987987987	10	35.0
	987987987	30	5.0
	987654321	30	20.0
	987654321	20	15.0
	888665555	20	null

PROJETO	PJNOME	PNUMERO	PLOCALIZACAO	DNUM
	ProdutoX	1	Bellaire	5
	ProdutoY	2	Sugarland	5
	ProdutoZ	3	Houston	5
	Automatização	10	Stafford	4
	Reorganização	20	Houston	1
	Novos Benefícios	30	Stafford	4

DEPENDENTE	ESSN	NOME_DEPENDENTE	SEXO	DATANASC	PARENTESCO
	333445555	Alice	F	1986-04-05	FILHA
	333445555	Theodore	M	1983-10-25	FILHO
	333445555	Joy	F	1958-05-03	CÔNJUGE
	987654321	Abner	M	1942-02-28	CÔNJUGE
	123456789	Michael	M	1988-01-04	FILHO
	123456789	Alice	F	1988-12-30	FILHA
	123456789	Elizabeth	F	1967-05-05	CÔNJUGE

a) Com a Álgebra Relacional descreva as seguintes consultas:

- Montar uma lista com o nome, sexo, data de nascimento e parentesco dos dependentes de cada empregado do sexo feminino, além do nome, SSN e Sexo do empregado.
- Mostre como ficaram as relações intermediárias.

b) Listar a SSN de todos os empregados do departamento 5 e os SSN's dos empregados que são supervisores diretos de algum empregado do departamento 5.

Junção-Theta (Junção)

Tabela_1 \bowtie condição Tabela_2

A operação Junção indicada por \bowtie é usada para combinar as tuplas relacionadas em duas relações dentro de uma tupla única.

A operação de Junção pode ser definida por um produto cartesiano seguido por uma seleção.

- Além da *igualdade* outras condições podem utilizadas em uma junção, como **menor** ($A < D$), conforme exercício abaixo.

Exercícios

1- Mostre o resultado da junção-theta aplicada sobre as relações U e V.

a)

A	B	C
1	2	3
6	7	8
9	7	8

B	C	D
2	3	4
2	3	5
7	8	10

U $\bowtie_{A < D}$ V					

b)

U $\bowtie_{A < D \text{ AND } U.B \neq V.B}$ V					

Para ilustrar a Junção, suponha que queiramos recuperar o nome do gerente de cada departamento. Para obter o nome do gerente precisaremos combinar cada tupla departamento com a tupla empregado na qual o valor do SSN case com o valor GERSSN da tupla departamento:

DEPT_GER \leftarrow DEPARTAMENTO \bowtie GERSSN = SSN EMPREGADO
RESULTADO $\leftarrow \pi_{\text{DNOME, UNOME, PNOME}}(\text{DEPTO_GER})$

As variações de Junção:

Equijoin:

- Quando a condição de uma junção é a igualdade, a junção é chamada de **Equijoin** (observar que a tabela resultante terá sempre duas colunas idênticas).

CIDADE $\bowtie_{\text{Cidade.UF} = \text{Estado.UF}}$ ESTADO

Cod_Cidade	Nome_Cidade	Capital	Cidade_UF	Estado_UF	Nome_Estado	Regiao
1234	Porto Alegre	S	RS	RS	Rio Grande do Sul	Sul
2345	São Paulo	S	SP	SP	São Paulo	Sudeste
1235	Caxias do Sul	N	RS	RS	Rio Grande do Sul	Sul

↑ ↑
colunas idênticas

Junção Natural

Tabela_1 ⋈ Tabela_2

Uma **Junção Natural** é um **Equijoin** em que uma das colunas idênticas é eliminada.

- O operador de junção natural combina as tuplas de duas relações que tem atributos comuns (mesmo nome), resultando numa relação que contém apenas as tuplas onde todos os atributos comuns apresentam o mesmo valor. A definição padrão de **Junção Natural** exige que os dois atributos de Junção tenham o mesmo nome em ambas as relações. Se esse não for o caso uma operação de **Rename** deve ser aplicada.
- Uma das colunas idênticas é eliminada, evitando a duplicidade:

Tabela_1

A	B	C	D	E
---	---	---	---	---

Tabela_2

D	E	F	G	H
---	---	---	---	---

Junção Natural

A	B	C	D	E	F	G	H
---	---	---	---	---	---	---	---

CIDADE ⋈ ESTADO

Cód_Cidade	Nome_Cidade	Capital	UF	Nome_Estado	Regiao
1234	Porto Alegre	S	RS	Rio Grande do Sul	Sul
2345	São Paulo	S	SP	São Paulo	Sudeste
1235	Caxias do Sul	N	RS	Rio Grande do Sul	Sul

Exercícios

1- Mostre o resultado da junção natural das relações U e V.

U	V			U ⋈ V
A	B	C	D	
1	2	3	4	
6	7	3	5	
9	7	8	10	

No exemplo seguinte, primeiro nós renomearemos o atributo DNUMERO de DEPARTAMENTO para DNUM -- Assim ele terá o mesmo nome do atributo DNUM em PROJETO -- então aplicaremos a Junção Natural:

$DEP_PROJ \leftarrow PROJETO * \rho_{(DNOME, DNUM, GERSSN, GERDATAINICIO)}(DEPARTAMENTO)$

A mesma consulta em dois passos:

$DEPT \leftarrow \rho_{(DNOME, DNUM, GERSSN, GERDATAINICIO)}(DEPARTAMENTO)$

$DEPT_PROJ \leftarrow PROJETO * DEPT$

Obs.: o atributo DNUM é chamado de atributo de Junção.

Inner Join

Originalmente o modelo relacional definiu o operador de junção apenas com as características apresentadas até aqui. Posteriormente, esse tipo de operador ficou conhecido como *Inner Join* e a operação de junção foi estendida para a definição de *Outer Join*.

Em um *Inner Join* existe a possibilidade de que algumas das linhas de uma ou ambas as tabelas de uma junção não façam parte da tabela resultante.

Exemplo:

FUNCIONARIO

Cod_Func	Nome	Cod_Proj
301	João	01
482	Miguel	02
127	Roberto	04
185	Dora	02
079	Esmeralda	01
246	Jaqueline	03
128	Leandro	--
248	Marcos	--

PROJETO

Cod_Proj	Projeto
01	Alfa Centauro
02	Betelgeuse
03	Canopus
04	Orion
05	Argos
06	Vênus

FUNCIONARIO ⋈ PROJETO

Cod_Func	Nome	Cod_Proj	Projeto
301	João	01	Alfa Centauro
482	Miguel	02	Betelgeuse
127	Roberto	04	Orion
185	Dora	02	Betelgeuse
079	Esmeralda	01	Alfa Centauro
246	Jaqueline	03	Canopus

De acordo com as regras de um *Inner Join*, as informações sobre os funcionários Leandro, Marcos e os projetos Argos e Vênus não foram incluídas na tabela resultante.

Outer Join

O *Outer Join* é uma extensão do operador de junção utilizado para tratamento de ausência de informações.

A operação *Outer Join* mantém na tabela resultante TODAS as linhas de uma ou mais tabelas participantes da junção.

- Utilizada quando é necessário incluir linhas de uma tabela onde os valores não têm correspondência com a outra.

Outer Join = Resultado do *Inner Join* +

Linhas da tabela sem correspondência (na outra tabela do *join*);

- As linhas da tabela que não tem correspondência são completadas com nulos.

Exemplo:

Linhas de FUNCIONARIO sem correspondência em PROJETO

Cod_Func	Nome	Cod_Proj	Projeto
128	Leandro	--	--
248	Marcos	--	--

Linhas de PROJETO sem correspondência em FUNCIONARIO

Cod_Func	Nome	Cod_Proj	Projeto
--	--	05	Argos
--	--	06	Vênus

Foram definidos três tipos de *Outer Join* entre duas tabelas:

- **Left Outer Join** 
- **Right Outer Join** 
- **Full Outer Join** 

Left Outer Join = *Inner Join* + linhas não participantes da tabela à esquerda

FUNCIONARIO  **PROJETO**

Cod_Func	Nome	Cod_Proj	Projeto
301	João	01	Alfa Centauro
482	Miguel	02	Betelgeuse
127	Roberto	04	Orion
185	Dora	02	Betelgeuse
079	Esmeralda	01	Alfa Centauro
246	Jaqueline	03	Canopus
128	Leandro	--	--
248	Marcos	--	--

Todas as linhas de FUNCIONARIO participam da tabela resultante

Right Outer Join = *Inner Join* + linhas não participantes da tabela à direita

FUNCIONARIO  **PROJETO**

Cod_Func	Nome	Cod_Proj	Projeto
301	João	01	Alfa Centauro
482	Miguel	02	Betelgeuse
127	Roberto	04	Orion
185	Dora	02	Betelgeuse
079	Esmeralda	01	Alfa Centauro
246	Jaqueline	03	Canopus
--	--	05	Argos
--	--	06	Vênus

Todas as linhas de PROJETO participam da tabela resultante

Full Outer Join = *Inner Join* + linhas não participantes de ambas tabelas

FUNCIONARIO  **PROJETO**

Cod_Func	Nome	Cod_Proj	Projeto
301	João	01	Alfa Centauro
482	Miguel	02	Betelgeuse
127	Roberto	04	Orion
185	Dora	02	Betelgeuse
079	Esmeralda	01	Alfa Centauro
246	Jaqueline	03	Canopus
128	Leandro	--	--
248	Marcos	--	--
--	--	05	Argos
--	--	06	Vênus

Todas as linhas de FUNCIONARIO e PROJETO participam da tabela resultante

Divisão

$$A \div B$$

Divisão é a operação entre duas tabelas onde:

- Uma é o dividendo e outra o divisor
- As tabelas devem possuir uma ou mais colunas em comum
- A tabela dividendo (A) deve possuir colunas adicionais
- A tabela resultante é formada pelas colunas da tabela dividendo (A) que não fazem parte da tabela divisora (B).
- A tabela resultante contém apenas linhas da tabela dividendo (A) que satisfazem a comparação com todas as linhas da tabela divisora (B)

DIVIDENDO

÷

DIVISOR

RESULTADO

Cod_Emp	Habilidade
301	73
301	02
301	11
482	02
482	11
127	73
127	43
127	02

Habilidade
02
11

Cod_Emp
301
482

Resultado: Lista de todos os empregados que possuem as habilidades '02' e '11'

• Operações Aninhadas

Podemos construir expressões complexas aplicando operadores sobre o resultado da aplicação de outros operadores e assim por diante

Exemplo: $\pi_{\text{Nome_Cidade, UF}}(\sigma_{\text{regiao='SUL'}}(CIDADE \bowtie ESTADO))$

$\pi_{\text{Nome_Cidade, UF}}(CIDADE \bowtie (\sigma_{\text{regiao='SUL'}}(ESTADO)))$

Renomeando uma relação

$$\rho_{S(A_1, A_2, \dots, A_n)}(R)$$

$$\rho_S(R)$$

Renomeia a relação "R" para "S", sem renomear os atributos.

$$\rho_{S(A_1, A_2, \dots, A_n)}(R)$$

Renomeia a relação "R" para "S", renomeando também os atributos resultantes para A_1, A_2, \dots, A_n na ordem da esquerda para a direita

A operação de renomear é útil quando necessitamos controlar os nomes dos atributos como, por exemplo, no produto cartesiano de uma relação com ela mesma.

Exercícios

Considerando as relações a seguir, escreva as expressões da álgebra relacional para responder as consultas, mostrando ainda o resultado das mesmas.

PRODUTO

fabricante	modelo	Tipo
A	1001	Pc
A	1002	Pc
A	1003	Pc
B	1004	Pc
B	1006	Pc
B	3002	impressora
B	3004	impressora
C	1005	Pc
C	1007	Pc
D	1008	Pc
D	1009	Pc
D	1010	Pc
D	2001	Laptop
D	2002	Laptop
D	2003	Laptop
D	3001	impressora
D	3003	impressora
E	2004	Laptop
E	2008	Laptop
F	2005	Laptop
G	2006	Laptop
G	2007	Laptop
H	3005	impressora
I	3006	impressora

PC

modelo	velocidade	memoria	hd	cd	preco
1001	133	16	1.6	6x	1595
1002	120	16	1.6	6x	1399
1003	166	24	2.5	6x	1899
1004	166	32	2.5	8x	1999
1005	166	16	2.0	8x	1999
1006	200	32	3.1	8x	2099
1007	200	32	3.2	8x	2349
1008	180	32	2.0	8x	3699
1009	200	32	2.5	8x	2599
1010	160	16	1.2	8x	1495

LAPTOP

modelo	velocidade	memoria	hd	tela	preco
2001	100	20	1.10	9.5	1999
2002	117	12	0.75	11.3	2499
2003	117	32	1.00	10.4	3599
2004	133	16	1.10	11.2	3499
2005	133	16	1.00	11.3	2599
2006	120	8	0.81	12.1	1999
2007	150	16	1.35	12.1	4799
2008	120	16	1.10	12.1	2099

IMPRESSORA

modelo	cor	tipo	preço		
3001	true	ink-jet	275		
3002	true	ink-jet	269		
3003	false	laser	829		
3004	false	laser	879		
3005	false	ink-jet	180		
3006	true	dry	470		

(fonte: A First Course in Database Systems. Ullmann & Widom. Prentice-Hall)

1. Quais os modelos de PC que têm velocidade maior ou igual a 150 ?
2. Que fabricantes fazem laptops com disco rígido maior ou igual a 1 GB ?
3. Encontre os números dos modelos e preços de todos os produtos (de qualquer tipo) produzidos pelo fabricante B.
4. Encontre os números dos modelos de todas as impressoras laser colorida.
5. Encontre todos os fabricantes que vendem laptops, mas não vendem PCs.
6. Encontre os tamanhos de discos rígidos que ocorrem em dois ou mais PCs.
7. Encontre os pares de modelos de PCs que apresentam a mesma velocidade e memória (um par somente deve aparecer uma vez, i.e., se deve listar (i,j) mas não (j,i)).
8. Encontre os fabricantes de ao menos dois computadores diferentes (PCs ou laptops) com velocidade maior ou igual a 133.
9. Encontre o(s) fabricante(s) do computador com maior velocidade disponível.
10. Encontre os fabricantes de PCs com ao menos três velocidades diferentes.
11. Encontre os fabricantes que vendem exatamente três modelos diferentes de PCs.