

Álgebra Relacional

- Um modelo de dados, além de definir estruturas e restrições, deve definir um conjunto de operações para manipular os dados.
 - A **álgebra relacional** constitui o conjunto básico de operações do modelo relacional.
 - As operações possibilitam especificar solicitações básicas de recuperação, sendo que o resultado de uma recuperação é uma nova relação, formada a partir de uma ou mais relações.
- As operações de recuperação se dividem em:
 - operações específicas de bancos de dados relacionais: seleção, projeção, junção, entre outras;
 - operações da teoria de conjuntos: união, interseção, diferença e produto cartesiano.

1

Empregado	PrimeiroNome	InicialMeio	UltimoNome	NumEmpregado	DataNascimento	Endereco	Sexo	Salario	NumSupervisor	NumDepto
João	B	Silva	123456789	09/01/65	R. da Bahia, 2557	M	300.00	333445555	5	
Frank	T	Santos	333445555	08/12/55	Av. Afonso Pena, 3005	M	4000.00	888665555	5	
Alice	J	Pereira	999887777	19/07/68	Av. do Contorno, 2534	F	2500.00	987654321	4	
Luciene	S	Ferreira	987654321	20/06/51	R. Iral, 175	F	430.00	888665555	4	
Pedro	K	Magalhães	666884444	15/09/52	Av. Silva Lobo, 2050	M	1200.00	333445555	5	
Daniela	A	Oliveira	453453453	31/07/62	R. Ataliba Lago, 250	F	2500.00	333445555	5	
Mateus	V	Mascarenhas	987987987	29/03/79	R. Conria, 12	M	2500.00	987654321	4	
Fábio	E	Lemos	888665555	10/11/47	R. Chile, 425	M	5500.00	null	1	

Departamento	NomeDepto	NumDepto	NumGerente	DataInicioGerencia
Pesquisa		5	333445555	22/05/98
Administração		4	987654321	01/01/95
Diretoria		1	888665555	19/06/01

Localizacao_Depto	NumDept	Localizacao	Projeto	NomeProj	NumProj	Localizacao	NumDepto
1	Savassi		Produto X	1	Burtis	5	
4	Centro		Produto Y	2	Pampulha	5	
5	Burtis		Produto Z	3	Contagem	5	
5	Pampulha		Informalização	10	Centro	4	
5	Contagem		Reorganização	20	Savassi	1	
			NovosBeneficios	30	Centro	4	

Trabalha_em	NumEmpregado	NumProj	Horas
	123456789	1	32
	123456789	2	7
	666884444	3	40
	453453453	1	20
	453453453	2	20
	333445555	2	10
	333445555	3	10
	333445555	10	10
	333445555	20	10
	999887777	30	30
	999887777	10	10
	987987987	10	35
	987987987	30	5
	987654321	30	20
	987654321	20	15
	888665555	20	null

Dependente	NumEmpregado	NomeDependente	Sexo	DataAniversario	Parentesco
	333445555	Aline	F	03/04/76	Filha
	333445555	Vitor	M	25/10/73	Filho
	333445555	Joana	F	03/05/98	Cônjuge
	987654321	Igor	M	29/02/52	Cônjuge
	123456789	Michel	M	01/01/88	Filho
	123456789	Aline	F	31/12/88	Filha
	123456789	Elizabeth	F	05/05/57	Cônjuge

Operação Seleção

- A operação **Seleção** é utilizada para selecionar um conjunto de tuplas de uma relação:
 - $\sigma_{\langle \text{cond} \rangle}(\langle R \rangle)$, onde $\langle \text{cond} \rangle$ é uma condição de seleção e $\langle R \rangle$ é o nome de uma relação.
- Ex.: selecionar todos os empregados que trabalham no departamento 5.

$\sigma_{\text{NumDepto}=5}(\text{Empregado})$

Primeiro Nome	Inicial Meio	UltimoNome	NumEmpregado	DataNascimento	Endereco	Sexo	Salario	NumSupervisor	NumDepto
João	B	Silva	123456789	09/01/65	R. da Bahia, 2557	M	300.00	333445555	5
Frank	T	Santos	333445555	08/12/55	Av. Afonso Pena, 3005	M	4000.00	888665555	5
Pedro	K	Magalhães	666884444	15/09/52	Av. Silva Lobo, 2050	M	1200.00	333445555	5
Daniela	A	Oliveira	453453453	31/07/62	R. Ataliba Lago, 250	F	2500.00	333445555	5

3

Operação Seleção

- É uma operação unária (feita em uma única relação).
- O grau (número de atributos) da relação resultante é o mesmo da relação original.
- A operação é comutativa:

$$\sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(\langle R \rangle)) = \sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond1} \rangle}(\langle R \rangle))$$
- Pode-se combinar uma cascata de operações Seleção em uma única operação Seleção:

$$\sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond3} \rangle}(\langle R \rangle))) = \sigma_{\langle \text{cond1} \rangle \text{E} \langle \text{cond2} \rangle \text{E} \langle \text{cond3} \rangle}(\langle R \rangle)$$

4

Operação Projeção

- A operação **Projeção** é utilizada para selecionar um conjunto de atributos de uma relação:

$\pi_{\langle \text{atributos} \rangle}(\langle R \rangle)$, onde $\langle \text{atributos} \rangle$ é uma lista de atributos dentre os atributos da relação R e $\langle R \rangle$ é o nome de uma relação.

- Ex.: listar o nome e o salário de todos os empregados.

$\pi_{\text{PrimeiroNome, UltimoNome, Salario}}(\text{Empregado})$

PrimeiroNome	UltimoNome	Salario
João	Silva	300.00
Frank	Santos	4000.00
Alice	Pereira	2500.00
Luciene	Ferreira	430.00
Pedro	Magalhães	1200.00
Daniela	Oliveira	2500.00
Mateus	Mascarenhas	2500.00
Fábio	Lemos	5500.00

5

Operação Projeção

- É uma operação unária (feita em uma única relação).
- Caso a lista de atributos inclua somente atributos que não sejam chaves de R, é possível que ocorram tuplas duplicadas.
 - A operação Projeção remove tuplas duplicadas de tal forma que o resultado seja uma relação válida.
 - Com isso, o número de tuplas na relação resultante é sempre menor ou igual ao número de tuplas da relação R.
- A operação não é comutativa; pode-se dizer que:

$$\pi_{\langle \text{lista1} \rangle}(\pi_{\langle \text{lista2} \rangle}(\langle R \rangle)) = \pi_{\langle \text{lista1} \rangle}(\langle R \rangle)$$

6

Sequência de Operações

- É comum aplicar diversas operações da álgebra relacional, uma após a outra (sequência de operações).
 - Pode-se escrever as operações na forma de uma única expressão ou aplicar uma operação a cada vez, criando relações de resultado intermediário; neste último caso, deve-se nomear as relações envolvidas.
- Ex.: listar o nome e o salário de todos os empregados que trabalham no departamento de número 5.

$\pi_{\text{PrimeiroNome, UltimoNome, Salario}}(\sigma_{\text{NumDepto} = 5}(\text{Empregado}))$

ou

$\text{Dep5_Emps} \leftarrow \sigma_{\text{NumDepto} = 5}(\text{Empregado})$

$\text{Resultado} \leftarrow \pi_{\text{PrimeiroNome, UltimoNome, Salario}}(\text{Dep5_Emps})$

7

Sequência de Operações

- Pode-se utilizar a técnica "sequência de operações" para renomear os atributos nas relações intermediárias e de resultado: basta listar os nomes dos novos atributos entre parênteses juntamente com os nomes das novas relações:

$\text{Dep5_Emps} \leftarrow \sigma_{\text{NumDepto} = 5}(\text{Empregado})$

$\text{Resultado}(\text{PNome, UNome, Sal}) \leftarrow$

$\pi_{\text{PrimeiroNome, UltimoNome, Salario}}(\text{Dep5_Emps})$

- Caso nenhuma renomeação seja aplicada em uma Seleção, os nomes dos atributos na relação resultante são os mesmos da relação original e estarão na mesma ordem.
- Para uma Projeção sem renomeação, a relação resultante possui os mesmos nomes dos atributos especificados na lista de projeção e aparecem na mesma ordem da lista.

8

Operação Renomeação

- A operação **Renomeação** é utilizada para renomear uma relação ou atributos da mesma:
 - $\rho_{S(b_1, b_2, \dots, b_n)}(\langle R \rangle)$ **ou** $\rho_S(\langle R \rangle)$ **ou** $\rho_{(b_1, b_2, \dots, b_n)}(\langle R \rangle)$, onde $\langle S \rangle$ é o novo nome da relação, $\langle b_1, b_2, \dots, b_n \rangle$ são os novos nomes dos atributos e $\langle R \rangle$ é a relação original. A primeira expressão renomeia tanto a relação quanto os atributos, a segunda renomeia apenas a relação e a terceira renomeia apenas os atributos.
- Ex.: listar o nome e o salário de todos os empregados que trabalham no departamento de número 5.

Dep5_Emps $\leftarrow \sigma_{\text{NumDepto}=5}(\text{Empregado})$

$\rho_{\text{Resultado}}(\text{PNome}, \text{UNome}, \text{Sal})$
 $(\pi_{\text{PrimeiroNome}, \text{UltimoNome}, \text{Salario}}(\text{Dep5_Emps}))$

9

Operações Teóricas de Conjuntos

- A álgebra relacional possui um grupo padrão de operações matemáticas sobre conjuntos.
 - As operações são binárias, ou seja, envolvem duas relações.
 - Para algumas operações, as relações devem possuir o mesmo tipo de tuplas, sendo consideradas compatíveis para união.
- Duas relações $R(A_1, A_2, \dots, A_n)$ e $S(B_1, B_2, \dots, B_n)$ são compatíveis para união se possuírem o mesmo grau "n" e se $\text{dom}(A_i) = \text{dom}(B_i)$ para $1 \leq i \leq n$.

10

Operações Teóricas de Conjuntos

- As operações teóricas de conjuntos que exigem relações compatíveis para união são:
 - União:** denotada por $R \cup S$, gera uma relação que inclui todas as tuplas que estão em R ou em S ou em ambas.
 - Interseção:** denotada por $R \cap S$, gera uma relação que inclui todas as tuplas que estão tanto em R quanto em S.
 - Diferença:** denotada por $R - S$, gera uma relação que inclui todas as tuplas que estão em R, mas não estão em S.
- A relação resultante das operações possui os mesmos nomes de atributos da primeira relação (R) envolvida nas operações.

11

Operações Teóricas de Conjuntos

- As operações de união e interseção são comutativas e associativas:

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

$$R \cup (S \cap T) = (R \cup S) \cap T \quad \text{e} \quad (R \cap S) \cup T = R \cap (S \cup T)$$
- Ex.: listar o número de todos os empregados que trabalham no departamento 5 ou que supervisionam diretamente um empregado que trabalhe no departamento 5.

Dep5_Emps $\leftarrow \sigma_{\text{NumDepto}=5}(\text{Empregado})$

Result1 $\leftarrow \pi_{\text{NumEmpregado}}(\text{Dep5_Emps})$

Result2 (NumEmpregado) $\leftarrow \pi_{\text{NumSupervisor}}(\text{Dep5_Emps})$

Resultado $\leftarrow \text{Result1} \cup \text{Result2}$

12

Operações Teóricas de Conjuntos

- A operação de conjunto binária **Produto Cartesiano**, representada por \times , é utilizada para combinar tuplas de duas relações de forma combinatória.
 - As relações não precisam ser compatíveis para união.
- O resultado de $R(A_1, A_2, \dots, A_n) \times S(B_1, B_2, \dots, B_m)$ é uma relação $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$, com "n+m" atributos.
 - A relação Q possui uma tupla para cada combinação de tuplas das relações envolvidas: se R possui n_R tuplas e S possui n_S tuplas, então Q possuirá $n_R * n_S$ tuplas.
- Não é uma operação muito usual pois gera tuplas que não fazem sentido.
 - Torna-se prática quando é seguida por uma Seleção que combina valores de atributos nas relações envolvidas.

13

Operações Teóricas de Conjuntos

- Ex.: listar, para cada empregado do sexo feminino, os nomes dos seus dependentes.


```
Emps_Mulheres ← σSexo='F' (Empregado)
Nomes_Emp (Nome, Sobrenome, NumEmp) ←
  πPrimeiroNome, UltimoNome, NumEmpregado (Emps_Mulheres)
Deps_Emp ← Nomes_Emp × Dependente
Deps_Certos ← σNumEmp = NumEmpregado (Deps_Emp)
Resultado ← πNome, Sobrenome, NomeDependente (Deps_Certos)
```
- Uma vez que a operação Produto Cartesiano, seguida da operação Seleção, é utilizada com frequência, foi definida uma operação especial, denominada **Junção**, para especificar tal sequência como uma única operação.

14

Operação Junção

- A operação **Junção** é utilizada para combinar tuplas relacionadas de duas relações em uma única tupla:

$$R \bowtie_{\langle \text{cond} \rangle} S$$
 onde R e S são relações e $\langle \text{cond} \rangle$ é uma condição de junção entre as relações.
- $R(A_1, A_2, \dots, A_n) \bowtie_{\langle \text{cond} \rangle} S(B_1, B_2, \dots, B_m)$ gera uma relação $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$, com "n+m" atributos.
 - A relação Q possui uma tupla para cada combinação de tuplas das relações envolvidas, sempre que a combinação satisfizer a condição de junção.
- Uma condição geral de junção é: $\langle \text{cond1} \rangle$ e $\langle \text{cond2} \rangle$ e ... e $\langle \text{condN} \rangle$, onde cada condição é da forma $A_i \theta B_j$:
 - A_i é atributo de R, B_j é atributo de S de mesmo domínio de A_i , e θ é um operador de comparação $\{=, <, \leq, >, \geq, \neq\}$.

15

Operação Junção

- Ex.: listar, para cada empregado do sexo feminino, os nomes dos seus dependentes.


```
Emps_Mulheres ← σSexo='F' (Empregado)
Nomes_Emp (Nome, Sobrenome, NumEmp) ←
  πPrimeiroNome, UltimoNome, NumEmpregado (Emps_Mulheres)
Deps_Certos ← Nomes_Emp ⋈NumEmp = NumEmpregado Dependente
Resultado ← πNome, Sobrenome, NomeDependente (Deps_Certos)
```
- Ex.: listar o nome do gerente de cada departamento.


```
Ger_Dep ← Departamento ⋈NumGerente = NumEmpregado Empregado
Resultado ← πNomeDepto, PrimeiroNome, UltimoNome (Ger_Dep)
```

16

Operação Junção

- A operação Junção mais comum, denominada **Equijunção**, envolve apenas condições de junção com comparações de igualdade.
 - Uma Equijunção onde os dois atributos da comparação têm o mesmo nome é chamada **Junção Natural**, sendo denotada por *; neste caso, apenas um dos atributos da comparação aparece na relação resultante e a condição de junção não é especificada.
- Ex.: listar, para cada empregado do sexo feminino, os nomes dos seus dependentes.

$\text{Emps_Mulheres} \leftarrow \sigma_{\text{Sexo}='F'}(\text{Empregado})$

$\text{Deps_Certos} \leftarrow \text{Emps_Mulheres} * \text{Dependente}$

$\text{Resultado} \leftarrow \pi_{\text{PrimeiroNome, UltimoNome, NomeDependente}}(\text{Deps_Certos})$

17

Operação Divisão

- A operação binária **Divisão**, representada por \div , é utilizada para um tipo especial de consulta que ocorre, algumas vezes, em aplicações de bancos de dados.
- A operação Divisão $R(Z) \div S(X)$ só pode ser aplicada se $X \subseteq Z$.
 - O resultado da divisão é uma relação T contendo o conjunto de atributos de R que não são atributos de S, ou seja, os atributos Z-X.
 - Uma tupla de T é formada por valores dos atributos Z-X de R cujos valores referentes dos atributos X de R combinaram com todos os valores dos atributos X de S.

18

Operação Divisão

- Ex.: listar o nome dos empregados que trabalham em todos os projetos nos quais "João Silva" trabalha.
 - 1º passo: recuperar os números de projetos nos quais "João Silva" trabalha.

$\text{Joao} \leftarrow \sigma_{\text{PrimeiroNome}='João' \text{ E } \text{UltimoNome}='Silva'}(\text{Empregado})$

$\text{Projs_Joao} \leftarrow \pi_{\text{NumProj}}(\text{Trabalha_em} * \text{Joao})$

- 2º passo: filtrar os atributos desejados da relação Trabalha_em.

$\text{NumEmps_NumProjs} \leftarrow \pi_{\text{NumEmpregado, NumProj}}(\text{Trabalha_em})$
- 3º passo: aplicar a Divisão entre as duas relações geradas; o resultado conterá os números dos empregados desejados.

$\text{NumEmps} \leftarrow \text{NumEmps_NumProjs} \div \text{Projs_Joao}$

$\text{Resultado} \leftarrow \pi_{\text{PrimeiroNome, UltimoNome}}(\text{NumEmps} * \text{Empregado})$

19

Funções de Agregação e Agrupamento

- Uma solicitação que pode ser expressa na álgebra relacional é a aplicação de **funções matemáticas de agregação** em coleção de valores do banco de dados.
 - As funções mais comuns aplicadas a coleção de valores numéricos são: **Sum** (soma), **Average** (média), **Maximum** (máximo), **Minimum** (mínimo), **Count** (contador de tuplas).
- Outra solicitação envolve o **agrupamento de tuplas** de uma relação por meio dos valores de alguns atributos e, logo após, a aplicação de uma função de agregação em cada grupo.
 - Um exemplo é agrupar as tuplas de empregados pelo "NumDepto": cada grupo é composto pelas tuplas de empregados que trabalham em um mesmo departamento.

20

Funções de Agregação e Agrupamento

- Uma função de agrupamento é definida da seguinte forma:

$\langle \text{atributos de agrupamento} \rangle \mathcal{F} \langle \text{funções de agregação} \rangle (\langle R \rangle)$

onde $\langle R \rangle$ é uma relação, $\langle \text{atributos de agrupamento} \rangle$ é uma lista de atributos de R responsável pelo agrupamento e $\langle \text{funções de agregação} \rangle$ é uma lista de pares de ($\langle \text{função} \rangle \langle \text{atributo} \rangle$): em cada um destes pares, $\langle \text{função} \rangle$ é uma das funções de agregação permitidas e $\langle \text{atributo} \rangle$ é um atributo de R cuja função de agregação será aplicada.

- A relação resultante possui os atributos de agrupamento e os resultados gerados pelas funções de agregação.
 - Haverá uma tupla para cada grupo gerado pelos atributos de agrupamento.

21

Funções de Agregação e Agrupamento

- Ex.: listar o nº de cada departamento, a quantidade de empregados em cada um e a média de seus salários.

$\rho_{\text{Resultado}} (\text{NumD}, \text{NumEmps}, \text{MediaSals})$
 $(\text{NumDepto} \mathcal{F} \text{Count NumEmpregado, Average Salario (Empregado)})$

- Caso não seja aplicada nenhuma renomeação, os atributos da relação resultante correspondentes às funções de agregação são, cada um deles, a concatenação do nome da função e o nome do atributo ($\langle \text{função} \rangle _ \langle \text{atributo} \rangle$).
 - No exemplo

$\text{NumDepto} \mathcal{F} \text{Count NumEmpregado, Average Salario (Empregado)}$
 os atributos da relação resultante são: NumDepto, Count_NumEmpregado, Average_Salario.

22

Funções de Agregação e Agrupamento

- Caso não seja especificado algum atributo de agrupamento, as funções de agregação são aplicadas em todas as tuplas da relação envolvida.
 - A relação resultante terá uma só tupla.

- O exemplo

$\mathcal{F} \text{Count NumEmpregado, Average Salario (Empregado)}$

recupera a quantidade total de empregados e a média de seus salários (atributos Count_NumEmpregado e Average_Salario).

23

Fechamento Recursivo

- Um tipo de operação que pode ser expresso na álgebra relacional é o **fechamento recursivo**.
 - Tal operação é aplicada a um auto-relacionamento entre tuplas do mesmo tipo.
- Ex.: listar todos os empregados que são supervisionados, em todos os níveis, pelo empregado "Fábio Lemos".

24

Fechamento Recursivo

- No nível 1, tem-se:

$$\text{Fabio_Num} \leftarrow \pi_{\text{NumEmpregado}}(\sigma_{\text{PrimeiroNome}='Fábio' \text{ E } \text{UltimoNome}='Lemos'}(\text{Empregado}))$$

$$\text{Supervisao}(\text{NumEmp}, \text{NumSup}) \leftarrow \pi_{\text{NumEmpregado}, \text{NumSupervisor}}(\text{Empregado})$$

$$\text{Resultado1}(\text{Num}) \leftarrow \pi_{\text{NumEmp}}(\text{Supervisao} \bowtie_{\text{NumSup} = \text{NumEmpregado}} \text{Fabio_Num})$$

- No nível 2, tem-se:

$$\text{Resultado2}(\text{Num}) \leftarrow \pi_{\text{NumEmp}}(\text{Supervisao} \bowtie_{\text{NumSup} = \text{Num}} \text{Resultado1})$$

- Para obter ambos os conjuntos de empregados supervisionados nos níveis 1 e 2 por "Fábio Lemos", aplica-se a união:

$$\text{Resultado} \leftarrow \text{Resultado1} \cup \text{Resultado2}$$

25

Operações de Junção Externa

- As operações de **Junção Externa** são extensões da Junção.
 - São utilizadas quando se deseja manter todas as tuplas de R, ou de S, ou de ambas as relações, no resultado da Junção, caso elas possuam **ou não** tuplas que se combinem nas relações.
- As operações classificam-se em:
 - Junção Externa à Esquerda: denotada por $R \sqsupset \bowtie S$, mantém todas as tuplas da relação R; se não há tupla de S que combine, os atributos de S são preenchidos com valores nulos.
 - Junção Externa à Direita: denotada por $R \bowtie \sqsubset S$, mantém todas as tuplas da relação S; se não há tupla de R que combine, os atributos de R são preenchidos com valores nulos.
 - Junção Externa Completa: denotada por $R \sqsupset \bowtie \sqsubset S$, mantém todas as tuplas em ambas as relações; se alguma tupla não combina, preenche os atributos referentes com nulos.

26

Operações de Junção Externa

- Ex.: listar o nome dos empregados e, se for o caso, o nome dos departamentos que eles gerenciam.

$$\text{Temp} \leftarrow \text{Empregado} \sqsupset \bowtie_{\text{NumEmpregado} = \text{NumGerente}} \text{Departamento}$$

$$\text{Resultado} \leftarrow \pi_{\text{PrimeiroNome}, \text{UltimoNome}, \text{NomeDepto}}(\text{Temp})$$

OU

$$\text{Temp} \leftarrow \text{Departamento} \bowtie \sqsubset_{\text{NumGerente} = \text{NumEmpregado}} \text{Empregado}$$

$$\text{Resultado} \leftarrow \pi_{\text{PrimeiroNome}, \text{UltimoNome}, \text{NomeDepto}}(\text{Temp})$$

27

Operação União Externa

- A operação **União Externa** serve para realizar a união entre tuplas de duas relações, caso as mesmas não sejam compatíveis para união.
 - Tal operação irá encontrar a união entre tuplas de duas relações que são parcialmente compatíveis, significando que apenas alguns atributos são compatíveis para união.
 - Os atributos que não são compatíveis para união, de qualquer relação, são mantidos na relação resultante e, caso não possuem valores para uma determinada tupla, seus valores são preenchidos com nulos.

28

Operação União Externa

- Por exemplo, uma União Externa pode ser aplicada entre as relações AlunoGrad (Nome, Dep, Período) e AlunoPos (Nome, Dep, Nível).
 - A relação resultante é R (Nome, Dep, Período, Nível).
 - Uma tupla em R proveniente de tuplas em ambas relações terá valores para todos os atributos.
 - Uma tupla em R proveniente apenas de uma tupla da relação AlunoGrad (aluno de graduação) terá valor nulo para o campo "Nível".
 - Uma tupla em R proveniente apenas de uma tupla da relação AlunoPos (aluno de pós-graduação) terá valor nulo para o campo "Período".