

Álgebra

- ❑ Na matemática, uma álgebra é um conjunto de *objetos* e um conjunto de *operações* sobre estes objetos
- ❑ Exemplo: *Aritmética*, conjunto de números e de operações sobre números (soma, subtração, ...)

Álgebra relacional

- ❑ Álgebra desenvolvida para descrever operações sobre uma base de dados relacional
- ❑ O conjunto de objetos são as *tabelas*:
- ❑ Uma operação possui como operandos e como resultado tabelas
- ❑ Porque aprender:
  - Compreendendo álgebra relacional é mais fácil apreender SQL
  - Não há SGBD que implementa álgebra diretamente como DML, mas SQL incorpora cada vez mais conceitos de álgebra
  - Algoritmos de otimização de consulta definidos sobre álgebra (possível uso internamente no SGBD)

Operadores da álgebra relacional

- ❑ Operadores sobre *conjuntos* (uma tabela é um conjunto de linhas):
  - União
  - Interseção
  - Diferença
  - Produto Cartesiano
- ❑ Operadores *específicos* da álgebra relacional:
  - Seleção
  - Projeção
  - Junção
  - Divisão
  - Renomeação

BD exemplo

Peça		NomePeça	CodPeça	PesoPeça	CidadePeça
P1	Eixo Rolamento Mancal	Eixo	Cinza	10	PoA
P2		Preto	16	Rio	SaoPaulo
P3		Verde	30		

Embarq		CodFornec	QtdeEmbarc
P1	F1		300
P1		F2	400
P2		F3	200
P2		F4	300
			350

Fornec		NomeFornec	StatusFornec	CidadeFornec
F1	5	Silva		SaoPaulo
F2		Souza	10	Rio
F3		Alvares	5	SaoPaulo
F4		Tavares	8	Rio

Operação de Seleção

- ❑ A Seleção tem como operando uma tabela. O resultado é uma tabela que contém as linhas que obedecem a um determinado critério.
- ❑ **Sintaxe:**  
 $\sigma <\text{critério de seleção}> (<\text{tabela}>)$   
onde <tabela> é o nome de uma tabela ou uma expressão de álgebra relacional que resulta em uma tabela e <critério de seleção> é uma expressão booleana que envolve literais e valores de atributos da tabela
- ❑ O resultado da seleção tem colunas com os mesmos nomes e domínios da tabela de entrada.

Exemplo de Seleção (BD de peças e fornecedores)

- $\sigma$  CodPeça = 'P1' (Peça)
- ❑ Resulta em uma tabela que contém todos os dados da peça de código P1.

CodPeça	NomePeça	CorPeça	PesoPeça	CidadePeça
P1	Eixo	Cinza	10	PoA

Exemplo de Seleção

- $\sigma$  (StatusFornec > 5 and CidadeFornec = 'Rio') (Fornec)
- ❑ Resulta em uma tabela com os dados de todas os fornecedores que tenham status maior que 5 e sejam do Rio.

CodFornec	NomeFornec	StatusFornec	CidadeFornec
F2	Souza	10	Rio
F4	Tavares	8	Rio

Exemplo de Seleção

- $\sigma$  (StatusFornec > 5) ( $\sigma$  (CidadeFornec = 'Rio') (Fornec))
- ❑ O mesmo que o anterior, mas usando duas expressões de Seleção *aninhadas*
- ❑ O critério de seleção envolve somente valores de atributos de uma linha. Não é possível especificar critérios que envolvam múltiplas linhas nem critérios que envolvem diferentes tabelas.

Operação de Projeção

❑ A *Projeção* tem como operando uma tabela. O resultado é uma tabela que contém apenas as colunas selecionadas.

❑ Sintaxe:

$\pi$  <lista de colunas> (<tabela>)

onde <tabela> é o nome de uma tabela ou uma expressão de álgebra relacional que resulta em uma tabela e <lista de colunas> é uma lista que contém nomes de colunas da tabela operando.

Exemplo de Projeção (BD de peças e fornecedores)

$\pi$  CodPeça, NomePeça (Peça)

❑ Resulta em uma tabela que contém os códigos e os nomes de todas as peças.

CodPeça	NomePeça
P1	Eixo
P2	Rolamento
P3	Mancal

Exemplo de Projeção

❑ A Projeção pode resultar também na *eliminação de linhas*, caso colunas que são parte da chave forem eliminadas.

- Uma tabela é um conjunto de linhas

Se uma coluna cujos valores distinguem diferentes linhas é eliminada, surgem linhas duplicadas na tabela, que devem ser eliminadas.

❑ Exemplo

$\pi$  CidadeFornec (Fornec)

Resulta em uma tabela que contém todas as cidades em que há fornecedores. Note-se que se houver múltiplos fornecedores na mesma cidade, as diferentes linhas são eliminadas.

CidadeFornec
SãoPaulo
Rio

Projeção generalizada

❑ De forma geral, expressões aritméticas podem ser usadas na lista de projeção:

$\pi$  CodPeça, NomePeça, PesoPeça \* 1.1322 (Peça)

❑ Problema:

- qual é o nome da terceira coluna?
- ver operador de renomeação adiante

Sequências de operadores

Operadores diferentes podem ser aninhados

$\pi$  CodFornec, QtdeEmbarc  
( $\sigma$  CodPeça = 'P1' (Embarq) )

Resulta em tabela com código de fornecedor e quantidade embarcada para cada embarque da peça de código P1.

CodFornec	QtdeEmbarc
F1	300
F2	400
F3	200

Operações da teoria de conjuntos

- A álgebra relacional empresta da teoria de conjuntos quatro operadores: União, Intersecção, Diferença e Produto Cartesiano
- Sintaxe da operação **União**:  
 $\langle \text{tabela} \rangle_1 \cup \langle \text{tabela} \rangle_2$
- Sintaxe da operação **Intersecção**:  
 $\langle \text{tabela} \rangle_1 \cap \langle \text{tabela} \rangle_2$
- Sintaxe da operação **Diferença**:  
 $\langle \text{tabela} \rangle_1 - \langle \text{tabela} \rangle_2$
- Nos três casos, a operação possui duas tabelas como operando. As tabelas devem ser compatíveis para união:
  - Possuir o mesmo número de colunas
  - O domínio da *i-ésima* coluna de uma tabela deve ser idêntico ao domínio da *i-ésima* coluna da outra.
- Quando os nomes das colunas forem diferentes, adota-se a convenção de usar os nomes das colunas da primeira tabela.

Exemplo de União

$\pi$  CodFornec (Embarq)  
 $\cap$   
 $\pi$  CodFornec  
( $\sigma$  StatusFornec > 5 (Fornec) )

- Obtém os códigos de todos os fornecedores que tem embarques e que tem status maior que 5

Operação Produto Cartesiano

- Sintaxe da operação Produto Cartesiano:  
 $\langle \text{tabela} \rangle_1 \times \langle \text{tabela} \rangle_2$
- O produto cartesiano possui como operandos duas tabelas.
- O resultado é uma tabela cujas linhas são a combinação das linhas das tabelas  $\langle \text{tabela} \rangle_1$  e  $\langle \text{tabela} \rangle_2$ , tomando-se uma linha da  $\langle \text{tabela} \rangle_1$  e concatenando a com uma linha da  $\langle \text{tabela} \rangle_2$ .
  - Total de colunas do produto cartesiano =  
Número de colunas da primeira tabela +  
Número de colunas da segunda tabela
  - Número de linhas do produto cartesiano =  
Número de linhas da primeira tabela x  
Número de linhas da segunda tabela.

Exemplo de Produto Cartesiano

Embarq X Peça

Embarq			Peça			
CodPeça	CodFornec	QtdeEmbarc	CodPeça	NomePeça	CorPeça	PesoPeça
P1	F1	300	P1	Eixo	Cinza	10
P1	F2	400	P1	Eixo	Cinza	10
P1	F3	200	P1	Eixo	Cinza	10
P2	F1	300	P1	Eixo	Cinza	10
P2	F4	350	P1	Eixo	Cinza	10
P1	F1	300	P2	Rolamento	Preto	16
P1	F2	400	P2	Rolamento	Preto	16
P1	F3	200	P2	Rolamento	Preto	16
P2	F1	300	P2	Rolamento	Preto	16
P2	F4	350	P2	Rolamento	Preto	16
P1	F1	300	P3	Mancal	Verde	30
P1	F2	400	P3	Mancal	Verde	30
P1	F3	200	P3	Mancal	Verde	30
P2	F1	300	P3	Mancal	Verde	30
P2	F4	350	P3	Mancal	Verde	30

A operação tem Produto Cartesiano não é usada isoladamente. Normalmente, ela é combinada com uma seleção que envolve as diversas tabelas multiplicadas.

Seleção combinada com Produto Cartesiano

$\pi$  NomePeça  
 $(\sigma \text{ Embarq.CodPeça=Peça.CodPeça}$   
 $(\text{Embarq X Peça}))$

Obtém os nomes de todas as peças para as quais há embarques

NomePeça
Eixo
Rolamento

Renomeação

- Operador para atribuir (dentro de uma consulta) um novo nome a uma tabela
- Sintaxe  
 $p \text{ <novo nome> (<nome de tabela>)}$
- A tabela denominada <nome de tabela> recebe a denominação <novo nome>
- Necessário quando, em uma consulta, é necessário acessar mais de uma linha da mesma tabela

Renomeação

- Tabela  
 $\text{EMP}(\text{CodEmp, NomeEmp, CodEmpGer})$   
 $\text{CodEmpGer referencia EMP}$
- Obter o nome de cada empregado, que tem gerente, seguido do nome de seu gerente  
 $\pi \text{ EMP.NomeEmp, EMPGER.NomeEmp}$   
 $(\sigma \text{ EMP.CodEmpGer = EMPGER.CodEmp}$   
 $(\text{EMP x}$   
 $\text{p EMPGER (EMP))})$

Renomeação(2)

Tabela

EMP(CodEmp, NomeEmp, CodEmpGer)  
CodEmpGer referencia EMP

Obter o nome de cada empregado seguido do nome do gerente de seu gerente , caso ele o possua

$\pi$  EMP.NomeEmp, EMPGERGER. NomeEmp  
( $\sigma$  EMP.CodEmpGer = EMPGER.CodEmp AND  
EMPGER.CodEmpGer = EMPGERGER.CodEmp  
(EMP x  
 $\rho$  EMPGER (EMP) x  
 $\rho$  EMPGERGER (EMP)))

Renomeação(2)

( $\pi$  EMP.NomeEmp, EMPGER. NomeEmp  
( $\sigma$  EMP.CodEmpGer = EMPGER.CodEmp  
(EMP x  
 $\rho$  EMPGER (EMP)))  
  
( $\pi$  EMP.NomeEmp, EMPGERGER. NomeEmp  
( $\sigma$  EMP.CodEmpGer = EMPGER.CodEmp AND  
EMPGER.CodEmpGer = EMPGERGER.CodEmp  
(EMP x  
 $\rho$  EMPGER (EMP) x  
 $\rho$  EMPGERGER (EMP)))

Renomeação

As colunas também podem ser renomeadas.  
Exemplo

$\rho$  EMPGER(CodEmpGer,  
NomeEmpGer,  
CodEmpGerGer)  
(EMP)

Operador de ponto fixo

- Obter o nome de cada empregado seguido do nome de seu gerente, bem como, obter o nome de cada empregado seguido do nome do gerente de seu gerente e assim por diante.
- Caso o número de níveis de gerência seja fixo, é possível resolver com uma série de uniões
- Caso o número de níveis de gerência seja variável, é necessária uma operação que implemente **recursividade**.
- Álgebra relacional original não implementa recursividade.
- Há extensões que implementam o operador de **ponto fixo** ( $\phi$ ) .

Operador de ponto fixo

Exemplo:

$$\begin{aligned} &\phi ( \text{CHEFIA=} \\ &\quad \pi \text{ EMP.CodEmp, EMP.CodEmpGer (EMP)} \\ &\quad \cup \\ &\quad \pi \text{ EMP.CodEmp, CHEFIA.CodEmpGer} \\ &\quad \quad (\sigma \text{ EMP.CodEmpGer} = \text{CHEFIA.CodEmp} \\ &\quad \quad \quad (\text{EMP x CHEFIA}) \\ &\quad \quad \quad ) \\ &\quad ) \end{aligned}$$

Operador de ponto fixo (avaliação)

$$\begin{aligned} &\phi ( \text{CHEFIA=} \\ &\quad ( \pi \text{ EMP.CodEmp, EMP.CodEmpGer (EMP)} \\ &\quad \cup \\ &\quad \pi \text{ EMP.CodEmp, CHEFIA.CodEmpGer} \\ &\quad \quad (\sigma \text{ EMP.CodEmpGer} = \text{CHEFIA.CodEmp} \\ &\quad \quad \quad (\text{EMP x CHEFIA}) \\ &\quad \quad \quad ) \\ &\quad ) \end{aligned}$$

Passo 1:  
CHEFIA inicia vazia;  
Relação CHEFIA é  
definida como igual  
a projeção de EMP

Operador de ponto fixo (avaliação)

$$\begin{aligned} &\phi ( \text{CHEFIA=} \\ &\quad ( \pi \text{ EMP.CodEmp, EMP.CodEmpGer (EMP)} \\ &\quad \cup \\ &\quad \pi \text{ EMP.CodEmp, CHEFIA.CodEmpGer} \\ &\quad \quad (\sigma \text{ EMP.CodEmpGer} = \text{CHEFIA.CodEmp} \\ &\quad \quad \quad (\text{EMP x CHEFIA}) \\ &\quad \quad \quad ) \\ &\quad ) \end{aligned}$$

Passo 1+i (i=0,...):  
Relação EMP é  
juntada com a  
relação CHEFIA  
definida no passo  
anterior

Operador de ponto fixo (avaliação)

$$\begin{aligned} &\phi ( \text{CHEFIA=} \\ &\quad ( \pi \text{ EMP.CodEmp, EMP.CodEmpGer (EMP)} \\ &\quad \cup \\ &\quad \pi \text{ EMP.CodEmp, CHEFIA.CodEmpGer} \\ &\quad \quad (\sigma \text{ EMP.CodEmpGer} = \text{CHEFIA.CodEmp} \\ &\quad \quad \quad (\text{EMP x CHEFIA}) \\ &\quad \quad \quad ) \\ &\quad ) \end{aligned}$$

Final:  
Operador termina  
quando o resultado  
não muda pela  
iteração

Operação de Junção

- ❑ A combinação de uma operação de seleção aplicada sobre uma operação de produto cartesiano é usual em aplicações de BD. É através dela que dados de tabelas relacionadas são associados. Por isso, foi criada a operação de junção que corresponde exatamente à sequência de operações em questão.

Sintaxe:

$\langle \text{tabela} \rangle_1 \bowtie \langle \text{critério} \rangle \langle \text{tabela} \rangle_2$   
onde  $\langle \text{tabela} \rangle$  é o nome de uma tabela ou uma expressão de álgebra relacional que resulta em uma tabela e  $\langle \text{critério} \rangle$  é uma expressão booleana envolvendo literais e valores de atributos das duas tabelas.

- ❑ A Junção tem como operandos duas tabela. O resultado é equivalente a executar:

$\sigma \quad \langle \text{critério} \rangle \quad (\langle \text{tabela} \rangle_1 \times \langle \text{tabela} \rangle_2)$

Junção theta, Equijunção e Junção natural

- ❑ Critério de junção:
  - qualquer expressão booleana, inclusive comparações do tipo  $<$ ,  $>$ ,  $<=$ ,  $>=$ , ... entre os valores de atributos das tabelas envolvidas na junção
  - Essa operação genérica de Junção é chamada de *Junção theta*.
- ❑ Na maior parte dos casos, o  $\langle \text{critério} \rangle$  de junção é uma expressão como mostrada no exemplo
  - Envolve apenas igualdade de valores de atributos de diferentes tabelas
  - Esse tipo de junção é chamada de *Equijunção* e representada pelo símbolo  $*$

Exemplo de Junção (BD de peças e fornecedores)

$\text{Embarq} \bowtie (\text{Embarq.CódFornec} = \text{Fornec.CódFornec}) \text{ Fornec}$

Associa cada linha de embarque com a correspondente linha de fornecedor.

Embarq			Fornec			
CodPeça	CodFornec	QtdeEmbarc	CodFornec	NomeFornec	StatusFornec	CidadeFornec
P1	F1	300	F1	Silva	5	SãoPaulo
P1	F2	400	F2	Souza	10	Rio
P1	F3	200	F3	Alvares	5	SãoPaulo
P2	F1	300	F1	Silva	5	SãoPaulo
P2	F4	350	F4	Tavares	8	Rio

Equijunção

- ❑ **Sintaxe da Equijunção:**  
 $\langle \text{tabela} \rangle_1 * (\langle \text{lista} \rangle_1), (\langle \text{lista} \rangle_2) \langle \text{tabela} \rangle_2$   
 $\langle \text{lista} \rangle_1$  e  $\langle \text{lista} \rangle_2$  são as listas dos nomes das colunas das tabelas 1 e 2 respectivamente cujos valores são comparados um a um, para fazer a junção
- ❑ A operação de Equijunção distingue-se da Junção theta pelo fato de eliminar a segunda coluna em cada um dos pares que são comparados (já que os valores da segunda coluna são idênticos aos primeiros).



Exemplo de Equijunção (BD de peças e fornecedores)

Embarq \* (CódFornec) , (CódFornec) Fornec

Associa cada linha de embarque com a correspondente linha de fornecedor.

Embarq		Fornec			
CodPeça	CodFornec	QtdeEmbarc	NomeFornec	StatusFornec	CidadeFornec
P1	F1	300	Silva	5	SãoPaulo
P1	F2	400	Souza	10	Rio
P1	F3	200	Alvares	5	SãoPaulo
P2	F1	300	Silva	5	SãoPaulo
P2	F4	350	Tavares	8	Rio

Junção Natural

- No caso acima, as colunas de junção possuem os mesmos nomes.
- Para estes casos existe a *Junção natural*, na qual as listas de nomes de colunas não necessitam ser especificadas.

Sintaxe da Junção natural:

<tabela><sub>1</sub> \* <tabela><sub>2</sub>

Exemplo de Junção natural

Associa cada linha de embarque com a correspondente linha de fornecedor.

Embarq \* Fornec

Divisão

- Como a Junção, a Divisão é uma operação de álgebra relacional que pode ser construída a partir de outras, e é útil para casos que aparecem frequentemente.

Sintaxe

<tabela><sub>1</sub> ÷ <tabela><sub>2</sub>

Semântica

A operação de divisão tem duas tabelas como operandos.

Os nomes das colunas e respectivos domínios da <tabela><sub>2</sub> (C2) devem estar contidos dentro dos nomes das colunas e respectivos domínios da <tabela><sub>1</sub> (C1).

A tabela resultante tem como nomes de colunas e domínios aqueles que aparecem na <tabela><sub>1</sub>, mas não aparecem na <tabela><sub>2</sub> (C1-C2). Para que uma linha apareça no resultado, é necessário que a sua concatenação com cada linha da <tabela><sub>2</sub> apareça também na <tabela><sub>1</sub>.

Exemplos de divisão

T1	CódPeça	CódFornec
	P1	F1
	P2	F1
	P3	F1
	P4	F1
	P5	F1
	P2	F2
	P4	F2
	P1	F3
	P2	F3
T2	CódPeça	CódFornec
	P1	P2
	P2	P3
	P3	P4
T1÷T2	CódFornec	CódFornec
	F1	F1
	F3	F2

Exemplo de Divisão (BD de peças e fornecedores)

$$(\pi \text{ CodFornec, CodPeça (Embarq)})$$
  
$$\div$$
  
$$(\pi \text{ CodPeça}$$
  
$$(\sigma \text{ CidadePeça='PoA' or CidadePeça='Rio'}$$
  
$$(\text{Peça})))$$

A consulta obtém os códigos dos fornecedores que possuem embarques para todas peças de 'PoA' ou 'Rio'

A palavra “todos” muitas vezes está associada à operação de divisão

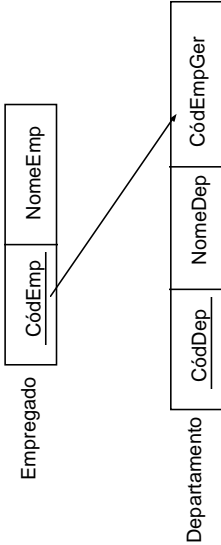
Conjunto mínimo de operações

- ☐ Muitas operações podem ser derivadas de outras
- ☐ Foi identificado um conjunto mínimo (completo) de operações, das quais todas as demais podem ser derivadas:
  - ☐ Seleção
  - ☐ Projeção,
  - ☐ União
  - ☐ Diferença
  - ☐ Produto Cartesiano

Restrições da operação de Junção

- ☐ A operação de junção concatena duas linhas das tabelas que estão sendo juntadas com base no critério de junção (normalmente por igualdade de valores de atributos)
- ☐ Uma linha que não possua nenhuma linha na outra tabela associada pelo critério de junção não aparece na tabela resultado.
- ☐ Há situações em que é necessário garantir que todas linhas de uma das tabelas de junção (ou de ambas) apareça no resultado.

Exemplo de restrição da junção



- ☐ Obter os dados de todos empregados junto com o nome de seu departamento, caso o empregado seja gerente do departamento
- ☐ Esta consulta não pode ser resolvida com a Junção, já que do resultado participariam apenas as linhas de empregados que são gerentes e não dos demais

Junção externa (outer join)

- Exemplo:
- Empregado

$\bowtie$  (CódEmp=CédEmpGer)

Departamento
- O operador  $\bowtie$  é chamado de *Junção externa esquerda* ("left outer-join")
- Semântica
- A Junção externa esquerda contém ao menos uma vez cada linha da tabela à esquerda do operador (no caso a tabela Empregado). Esta linha aparece concatenada com uma linha vazia, caso o critério de junção não seja verdadeiro para nenhuma linha da tabela à direita do operador de junção. Caso o critério de junção seja verdadeiro para uma ou mais linhas da tabela à direita, a linha da tabela à esquerda aparecerá concatenada com uma ou mais linhas da tabela à direita.
- De forma similar podem ser definidas:
- Junção externa direita (símbolo  $\bowtie\leftarrow$ )
- Junção externa plena (símbolo  $\bowtie\leftarrow\rightarrow$ )

Junção externa sem uso de sintaxe explícita

- Junção externa pode ser representada usando as junções internas

Empregado

$\bowtie$  (CódEmp=CédEmpGer)

Departamento

P

CodEmp, NomeEmp, NULL, NULL, NULL

(

$\pi$

CodEmp (Empregado)

-

$\pi$

CodEmpGer (Departamento)

)

\*

Empregado