

Álgebra relacional – junção e divisão

Carlos A. Heuser
2005

06/1

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 seqüência 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:

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

06/2

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	Álvares	5	SãoPaulo
P2	F1	300	F1	Silva	5	SãoPaulo
P2	F4	350	F4	Tavares	8	Rio

06/3

Junção theta, Equi-junção 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 **Equi-junção**

06/4

Equi-junção

❑ Sintaxe da Equi-junção:

$\langle \text{tabela} \rangle_1 \bowtie (\langle \text{lista} \rangle_1), (\langle \text{lista} \rangle_2) \langle \text{tabela} \rangle_2$

$\langle \text{lista} \rangle_1$ 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 Equi-junção 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).

06/5

Exemplo de Equijunção (BD de peças e fornecedores)

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

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	Álvares	5	SãoPaulo
P2	F1	300	Silva	5	SãoPaulo
P2	F4	350	Tavares	8	Rio

Coluna CodFornec pode ser referenciada por:

CodFornec

Embarq.CodFornec

Fornec.CodFornec

06/6

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:

$\langle \text{tabela} \rangle_1 \bowtie \langle \text{tabela} \rangle_2$

❑ Exemplo de Junção natural

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

$\text{Embarq} \bowtie \text{Fornec}$

06/7

Exercício – reescrever a consulta abaixo com os vários tipos de junção

❑ Tabela

$\text{EMP}(\underline{\text{CodEmp}}, \text{NomeEmp}, \text{CodEmpGer})$

CodEmpGer referencia EMP

- ❑ Obter o nome de cada empregado, que tem gerente, seguido do nome de seu gerente

$\pi \text{EMP.NomeEmp}, \text{EMPGER.NomeEmp}$

$(\sigma \text{EMP.CodEmpGer} = \text{EMPGER.CodEmp})$

$(\text{EMP} \times \rho \text{EMPGER}(\text{EMP}))$

06/8

Exercício – reescrever a consulta abaixo com os vários tipos de junção

- Obter o nome de cada empregado, que tem gerente, seguido do nome de seu gerente

$$\pi \text{ EMP.NomeEmp, EMPGER. NomeEmp}$$

$$(\sigma \text{ EMP.CodEmpGer} = \text{EMPGER.CodEmp}$$

$$(\text{EMP} \times \rho \text{ EMPGER} (\text{EMP})))$$

- Theta-join:**

$$\pi \text{ EMP.NomeEmp, EMPGER. NomeEmp}$$

$$(\text{EMP} \bowtie (\text{EMP.CodEmpGer} = \text{EMPGER.CodEmp})$$

$$\rho \text{ EMPGER} (\text{EMP}))$$

- Equi-join:**

$$\pi \text{ EMP.NomeEmp, EMPGER. NomeEmp}$$

$$(\text{EMP} \bowtie (\text{CodEmpGer}) (\text{CodEmp})$$

$$\rho \text{ EMPGER} (\text{EMP}))$$

06/9

Restrições da operação de Junção (interna)

- 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.

06/10

Exemplo

- Obter o nome de cada fornecedor, seguido do código de cada peça embarcada pelo fornecedor.

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

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

Fornec			
CodFornec	NomeFornec	StatusFornec	CidadeFornec
F1	Silva	5	SãoPaulo
F2	Souza	10	Rio
F3	Álvares	5	SãoPaulo
F4	Tavares	8	Rio

06/11

Exemplo

Obter o nome de cada fornecedor, seguido do código de cada peça embarcada pelo fornecedor. Caso o fornecedor, não embarque peças, seu nome deve aparecer seguido de vazio.

$$(\pi \text{ NomeFornec, CodPeca}$$

$$(\text{Fornec} \bowtie \text{Embarq}))$$

$$\cup$$

$$(\pi \text{ NomeFornec, NULL} (\text{Fornec})$$

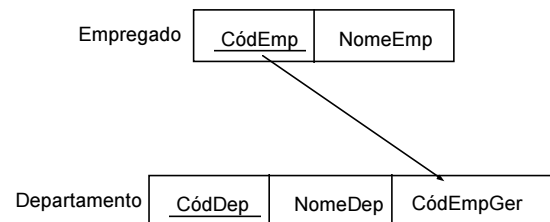
$$-$$

$$\pi \text{ NomeFornec, NULL} (\text{Embarq} \bowtie \text{Fornec}))$$

- Esta operação chama-se **junção externa (outer-join)**

06/12

Exemplo de restrição da junção

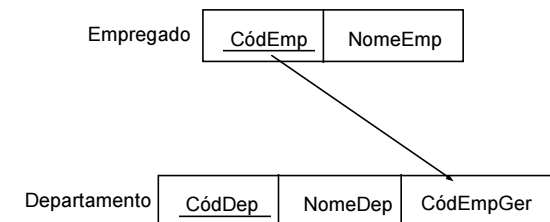


- Obter os dados de todos empregados junto com o nome de seu departamento, caso o empregado seja gerente do departamento

`Empregado ⋈ (CodEmp) (CodEmpGer) Departamento`

06/13

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

06/14

Junção externa (outer join)

- Exemplo:

`Empregado ⋈ (CódEmp=CodEmpGer) 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 \ltimes)
- **Junção externa plena** (símbolo $\ltimes\bowtie$)

06/15

Junção interna

A	B		B	C		A	B	C
1	X		X	10		1	X	10
2	Y		X	11		1	X	11
3	Z		Y	10		2	y	10
4	W		T	12		3	z	13
5	X		R	11		5	X	10
6	Y		S	12		5	X	11
7	Z		Z	13		6	Y	10
					=	7	Z	13

06/16

Junção interna

A	B
1	X
2	Y
3	Z
4	W
5	X
6	Y
7	Z



B	C
X	10
X	11
Y	10
T	12
R	11
S	12
Z	13

=

A	B	C
1	X	10
1	X	11
2	Y	10
3	Z	13
5	X	10
5	X	11
6	Y	10
7	Z	13

Linhas que não casam não aparecem no resultado

06/17

Junção externa à esquerda

A	B
1	X
2	Y
3	Z
4	W
5	X
6	Y
7	Z



B	C
X	10
X	11
Y	10
T	12
R	11
S	12
Z	13

=

A	B	C
1	X	10
1	X	11
2	Y	10
3	Z	13
4	W	NULL
5	X	10
5	X	11
6	Y	10
7	Z	13

06/18

Junção externa à direita

A	B
1	X
2	Y
3	Z
4	W
5	X
6	Y
7	Z



B	C
X	10
X	11
Y	10
T	12
R	11
S	12
Z	13

=

A	B	C
1	X	10
1	X	10
2	Y	10
3	Z	13
5	X	10
5	X	11
6	Y	10
NULL	T	12
NULL	R	11
NULL	S	12
7	Z	13

06/19

Junção externa total

A	B
1	X
2	Y
3	Z
4	W
5	X
6	Y
7	Z



B	C
X	10
X	11
Y	10
T	12
R	11
S	12
Z	13

=

A	B	C
1	X	10
1	X	10
2	Y	10
3	Z	13
4	W	NULL
5	X	10
5	X	11
6	Y	10
NULL	T	12
NULL	R	11
NULL	S	12
7	Z	13

06/20

Combinações de junções

Os três tipos de junções:

(theta, equi e natural)

podem ser combinados com as quatro alternativas:

(junção interna, junção externa à esquerda, junção externa à direita e junção externa total)

formando 12 combinações de junção.

	Interna	Externa à esquerda	Externa à direita	Externa total
Junção theta	\bowtie (<condição>)	$\leftarrow \bowtie$ (<condição>)	$\bowtie \rightarrow$ (<condição>)	\bowtie (<condição>)
Equi-junção	\bowtie ($c_1, c_{2,}, (c_1, c_{2,})$)	$\leftarrow \bowtie$ ($c_1, c_{2,}, (c_1, c_{2,})$)	$\bowtie \rightarrow$ ($c_1, c_{2,}, (c_1, c_{2,})$)	\bowtie ($c_1, c_{2,}, (c_1, c_{2,})$)
Junção natural	\bowtie	$\leftarrow \bowtie$	$\bowtie \rightarrow$	\bowtie

06/21

Exemplo

Obter o nome de cada fornecedor, seguido do código de cada peça embarcada pelo fornecedor. Caso o fornecedor, não embarque peças, seu nome deve aparecer seguido de vazio.

```
( $\pi$  NomeFornec, CodPeca
  (Fornec  $\leftarrow \bowtie$  Embarq))
```

06/22

Exercício – reescrever a consulta abaixo com os vários tipos de junção

Tabela

```
EMP(CodEmp, NomeEmp, CodEmpGer)
CodEmpGer referencia EMP
```

Obter o nome de cada empregado, que tem gerente, seguido do nome de seu gerente. Se o empregado não tiver gerente, seu nome deve aparecer seguido de vazio.

```
 $\pi$  EMP.NomeEmp, EMPGER. NomeEmp
  (EMP  $\leftarrow \bowtie$  (CodEmpGer) (CodEmp)
     $\rho$  EMPGER (EMP))
```

06/23

Junção externa sem uso de sintaxe explícita

Junção externa pode ser representada usando as junções internas

```
Empregado  $\bowtie$  (CodEmp=CodEmpGer) Departamento
 $\cup$ 
 $\pi$  CodEmp, NomeEmp, NULL, NULL, NULL
  (
     $\pi$  CodEmp (Empregado)
    -
     $\pi$  CodEmpGer (Departamento)
  )
 $\bowtie$  Empregado
```

06/24

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

$\langle \text{tabela} \rangle_1 \div \langle \text{tabela} \rangle_2$

Semântica

- A operação de divisão tem duas tabelas como operandos.
- Os nomes das colunas e respectivos domínios da $\langle \text{tabela} \rangle_2$ (C2) devem estar contidos dentro dos nomes das colunas e respectivos domínios da $\langle \text{tabela} \rangle_1$ (C1).
- A tabela resultante tem como nomes de colunas e domínios aqueles que aparecem na $\langle \text{tabela} \rangle_1$, mas não aparecem na $\langle \text{tabela} \rangle_2$ (C1-C2). Para que uma linha apareça no resultado, é necessário que a sua concatenação com cada linha da $\langle \text{tabela} \rangle_2$ apareça também na $\langle \text{tabela} \rangle_1$.

06/25

Exemplos de divisão

T1	CódPeça	CódFornec
	P1	F1
	P2	F1
	P3	F1
	P4	F1
	P5	F1
	P2	F2
	P4	F2
	P1	F3
	P2	F3
	P3	F3
	P4	F3
	P1	F4
	P2	F4
	P2	F5
T2	CódPeça	CódPeça
	P1	P2
	P2	P4
	P3	
T1÷T2	CódFornec	CódFornec
	F1	F1
	F3	F3

06/26

Exemplo de Divisão (BD de peças e fornecedores)

$(\pi_{\text{CodFornec, CodPeça}} (\text{Embarq}))$

\div

$(\pi_{\text{CodPeça}} (\sigma_{\text{CidadePeça}='PoA' \text{ or } \text{CidadePeça}='Rio'} (\text{Peça})))$

- 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.

06/27

Divisão "a pé"

- O operador de divisão pode ser definido a partir de outros (fornecedores com embarques de todas peças)

$\pi_{\text{CodFornec}} (\text{Embarq})$

$-$

$\pi_{\text{CodFornec}} ($
 $(\pi_{\text{CodFornec}} (\text{Embarq}) \times \pi_{\text{CodPeca}} (\text{Peca}))$
 $-$
 $\pi_{\text{CodFornec, CodPeca}} (\text{Embarq})$
 $)$

06/28

Divisão "a pé"

- O operador de divisão pode ser definido a partir de outros (fornecedores com embarques de todas peças)

```

π CodFornec (Embarq)
-
π CodFornec (
  ( π CodFornec (Embarq) x π CodPeca (Peca))
  -
  π CodFornec,CodPeca (Embarq)
)
    
```

Aqui cada fornecedor de embarque é casado com cada peça

06/29

Divisão "a pé"

- O operador de divisão pode ser definido a partir de outros (fornecedores com embarques de todas peças)

```

π CodFornec (Embarq)
-
π CodFornec (
  ( π CodFornec (Embarq) x π CodPeca (Peca))
  -
  π CodFornec,CodPeca (Embarq)
)
    
```

Aqui são retirados os pares <fornecedor,embarque> que existem em embarque

06/30

Divisão "a pé"

- O operador de divisão pode ser definido a partir de outros (fornecedores com embarques de todas peças)

```

π CodFornec (Embarq)
-
π CodFornec (
  ( π CodFornec (Embarq) x π CodPeca (Peca))
  -
  π CodFornec,CodPeca (Embarq)
)
    
```

O resultado são pares <fornecedor,embarque> onde o fornecedor **não embarcou** a peça. Fornecedores que embarcaram todas peças **não** aparecem

06/31

Divisão "a pé"

- O operador de divisão pode ser definido a partir de outros (fornecedores com embarques de todas peças)

```

π CodFornec (Embarq)
-
π CodFornec (
  ( π CodFornec (Embarq) x π CodPeca (Peca))
  -
  π CodFornec,CodPeca (Embarq)
)
    
```

assim resultam os códigos dos fornecedores que **não embarcaram** pelo menos uma das peças

06/32

Divisão "a pé"

- (fornecedores com embarques de todas peças)

(todos fornecedores)

$\pi \text{ CodFornec (Embarq)}$

-

$\pi \text{ CodFornec (}$
 $\quad (\pi \text{ CodFornec (Embarq)} \times \pi \text{ CodPeca (Peca))}$
 $\quad -$
 $\quad \pi \text{ CodFornec, CodPeca (Embarq)}$
 $\left. \vphantom{\pi \text{ CodFornec (Embarq)}} \right)$

menos

(fornecedores que não embarcaram uma das peças)

=

fornecedores que embarcaram todas peças

06/33

Conjunto mínimo de operações

- Muitas operações podem ser derivadas de outras.
- Foi identificado um **conjunto mínimo** (completo) de operações, das quais todas as demais operações da álgebra original podem ser derivadas:
 - Seleção
 - Projeção,
 - União
 - Diferença
 - Produto Cartesiano
- Observar que a operação de ponto fixo (ver adiante) não faz parte da álgebra relacional original e não é derivável deste conjunto de operações.

06/34

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** (ϕ).

06/35

Operador de ponto fixo

- Exemplo:

$\phi ($
 $\quad \text{CHEFIA} =$
 $\quad ($
 $\quad \quad \pi \text{ EMP.CodEmp, EMP.CodEmpGer (EMP)}$
 $\quad \cup$
 $\quad \quad \pi \text{ EMP.CodEmp, CHEFIA.CodEmpGer}$
 $\quad \quad \quad (\sigma \text{ EMP.CodEmpGer} = \text{CHEFIA.CodEmp}$
 $\quad \quad \quad \quad (\text{EMP} \times \text{CHEFIA})$
 $\quad \quad)$
 $\quad)$
 $)$

06/36

Operador de ponto fixo (avaliação)

```
φ (
  CHEFIA=
  (
    π EMP.CodEmp, EMP.CodEmpGer (EMP)
  ∪
    π EMP.CodEmp, CHEFIA.CodEmpGer
    (σ EMP.CodEmpGer = CHEFIA.CodEmp
    (EMP x CHEFIA)
  )
)
```

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

06/37

Operador de ponto fixo (avaliação)

```
φ (
  CHEFIA=
  (
    π EMP.CodEmp, EMP.CodEmpGer (EMP)
  ∪
    π EMP.CodEmp, CHEFIA.CodEmpGer
    (σ EMP.CodEmpGer = CHEFIA.CodEmp
    (EMP x CHEFIA)
  )
)
```

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

06/38

Operador de ponto fixo (avaliação)

```
φ (
  CHEFIA=
  (
    π EMP.CodEmp, EMP.CodEmpGer (EMP)
  ∪
    π EMP.CodEmp, CHEFIA.CodEmpGer
    (σ EMP.CodEmpGer = CHEFIA.CodEmp
    (EMP x CHEFIA)
  )
)
```

Final:
Operador termina
quando o resultado
não muda pela
iteração

06/39