

[MAC0313]
Introdução aos Sistemas de Bancos de Dados
Aula 9
A Álgebra Relacional (Continuação)

Kelly Rosa Braghetto

DCC-IME-USP

3 de setembro de 2014

Operações da teoria dos conjuntos

Produto Cartesiano (ou Produto Cruzado)

- ▶ É uma operação binária
- ▶ **Não requer compatibilidade de união** entre as relações envolvidas
- ▶ O produto cartesiano entre duas relações $R(A_1, A_2, \dots, A_n)$ e $S(B_1, B_2, \dots, B_m)$ é denotado por

$$R \times S$$

- ▶ Resultado: relação Q com as tuplas formadas a partir da combinação de cada uma das tuplas em R com cada uma das tuplas em S
- ▶ Q tem $n + m$ atributos; $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$
- ▶ Se $|R| = n_r$ e $|S| = n_s$, então $|Q| = n_r \times n_s$

Operações da teoria dos conjuntos

Produto Cartesiano (ou Produto Cruzado)

A	B
1	2
3	4

Relação R

B	C	D
2	5	6
4	7	8
9	10	11

Relação S

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

Resultado $R \times S$

Operações da teoria dos conjuntos

Produto Cartesiano – com renomeação antes do produto

A	B
1	2
3	4

Relação R

B	C	D
2	5	6
4	7	8
9	10	11

Relação S

A	B	X	C	D
1	2	2	5	6
1	2	4	7	8
1	2	9	10	11
3	4	2	5	6
3	4	4	7	8
3	4	9	10	11

Resultado $R \times \rho_{S(X,C,D)}(S)$

Operação Produto Cartesiano (ou Produto Cruzado)

- ▶ A operação é particularmente útil quando seguida de uma operação de seleção, para recuperar informações sobre entidades que se relacionam no BD
- ▶ Ex: Listar o nome dos dependentes de cada funcionária

$$\text{FUNC_MULHERES} \leftarrow \sigma_{\text{Sexo}='F'}(\text{FUNCIONARIO})$$

$$\text{FUNC_NOMES} \leftarrow \pi_{Pnome, Unome, Cpf}(\text{FUNC_MULHERES})$$

$$\text{FUNC_DEPENDENTES} \leftarrow \text{FUNC_NOMES} \times \text{DEPENDENTES}$$

$$\text{DEPENDENTE_PARTIC} \leftarrow \sigma_{\text{Cpf}=F\text{cpf}}(\text{FUNC_DEPENDENTES})$$

$$\text{RESULTADO} \leftarrow \pi_{Pnome, Unome, Nome_{dependente}}(\text{DEPENDENTE_PARTIC})$$

FUNCIONARIO

Pnome	Minicial	Unome	<u>Cpf</u>	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
-------	----------	-------	------------	----------	----------	------	---------	----------------	-----

DEPENDENTE

<u>Fopf</u>	<u>Nome_dependente</u>	Sexo	Datanasc	Parentesco
-------------	------------------------	------	----------	------------

Operações relacionais binárias

Junção (ou Junção Theta)

- ▶ É usada para **combinar duplas relacionadas** de duas relações
- ▶ Operação muito importante em BDs relacionais – possibilita processar relacionamentos

Operações relacionais binárias

Junção (ou Junção Theta)

- ▶ Forma geral da junção sobre duas relações $R(A_1, A_2, \dots, A_n)$ e $S(B_1, B_2, \dots, B_m)$:

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

- ▶ Resultado: relação Q com $n + m$ atributos;
 $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$
- ▶ Para toda tupla t resultante da combinação de uma tupla de R com uma de S , se t satisfaz a condição de junção, então $t \in Q$
- ▶ A **<condição de junção>** é definida sobre os atributos de R e S e é avaliada para todas as combinações de tuplas possíveis

Operações relacionais binárias

Junção (ou Junção Theta)

- ▶ Forma geral da <condição de junção>:

<condição1> AND <condição2> AND ... AND <condição p >

onde cada <condição k > tem a forma A_i **op** B_j e

- ▶ A_i é um atributo de R
- ▶ B_j é um atributo de S
- ▶ A_i e B_j têm o mesmo domínio
- ▶ **op** é um dos operadores de comparação ($=, <, \leq, >, \geq$ ou \neq)

Operações relacionais binárias

Junção (ou Junção Theta)

A	B
1	2
3	4

Relação R

B	C	D
2	5	6
4	7	8
9	10	11

Relação S

A	R.B	S.B	C	D
1	2	2	5	6
3	4	4	7	8

Resultado $R \bowtie_{R.B=S.B} S$

A	R.B	S.B	C	D
1	2	4	7	8
1	2	9	10	11
3	4	9	10	11

Resultado $R \bowtie_{R.B < S.B} S$

Operações relacionais binárias

FUNCIONARIO

Pnome	Minicial	Unome	<u>Cpf</u>	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
-------	----------	-------	------------	----------	----------	------	---------	----------------	-----

DEPARTAMENTO

Dnome	<u>Dnumero</u>	Cpf_gerente	Data_inicio_gerente
-------	----------------	-------------	---------------------

Junção (ou Junção Theta)

- ▶ Ex.: Recuperar o nome do gerente de cada departamento.

$$\text{DEP_GER} \leftarrow \text{DEPARTAMENTO} \bowtie_{\text{Cpf_gerente} = \text{Cpf}} \text{FUNCIONARIO}$$

$$\text{RESULTADO} \leftarrow \pi_{Dnome, Pnome, Unome}(\text{DEP_GER})$$

Operações relacionais binárias

Junção (ou Junção Theta)

Não aparecem no resultado de uma junção:

- ▶ Tuplas cujos valores dos atributos não satisfazem a condição de junção
- ▶ Tuplas que possuem NULL como valor para algum *atributo de junção* (ou seja, para um atributo que aparece na condição de junção)

Se nenhuma combinação de tuplas satisfizer a condição de junção, o resultado da operação será uma relação vazia.

Operações relacionais binárias

Variações de Junção

- ▶ **Equijunção** – a condição de junção envolve apenas comparações de igualdade
 - ▶ As tuplas do resultado sempre contêm 1 ou mais pares de atributos com valores idênticos
- ▶ **Junção natural** (expressa pelo operador $*$) – equivale a uma equijunção, seguida da remoção dos atributos desnecessários
 - ▶ Essa operação **requer que cada par de atributos de junção tenham o mesmo nome nas duas relações**; se isso não acontecer, é preciso renomear os atributos de uma das relações antes de aplicar a junção natural

Operações relacionais binárias

Equijunção e Junção natural

A	B
1	2
3	4

Relação R

B	C	D
2	5	6
4	7	8
9	10	11

Relação S

A	R.B	S.B	C	D
1	2	2	5	6
3	4	4	7	8

Resultado da equijunção

$$R \bowtie_{R.B=S.B} S$$

A	B	C	D
1	2	5	6
3	4	7	8

Resultado da junção natural

$$R * S$$

Operações relacionais binárias

DEPARTAMENTO

Dnome	<u>Dnumero</u>	Cpf_gerente	Data_inicio_gerente
-------	----------------	-------------	---------------------

PROJETO

Projnome	<u>Projnumero</u>	Projlocal	Dnum
----------	-------------------	-----------	------

Junção Natural

- ▶ Ex.: Combinar cada tupla de PROJETO com a tupla de DEPARTAMENTO que controla o projeto

PROJETO_DEP \leftarrow PROJETO *

$\rho(Dnome, Dnum, Cpf_gerente, Data_inicio_gerente)(DEPARTAMENTO)$

Conjunto completo de operações da álgebra relacional

$$\{\sigma, \pi, \cup, \rho, -, \times\}$$

- ▶ Qualquer uma das outras operações relacionais podem ser expressas como uma sequência de operações do conjunto acima
- ▶ Exemplos:
 - ▶ Intersecção: $R \cap S \equiv (R \cup S) - ((R - S) \cup (S - R))$
 - ▶ Junção: $R \bowtie_{\langle \text{condição} \rangle} \equiv \sigma_{\langle \text{condição} \rangle}(R \times S)$

A operação de Divisão (\div)

- ▶ Foi inserida na álgebra relacional básica por conveniência (não por necessidade!)
- ▶ A operação $R \div S$ só se aplica a duas relações $R(Z)$ e $S(X)$ em que os atributos de S são um subconjunto dos atributos de R , ou seja, $X \subseteq Z$
- ▶ O resultado de $R \div S$ é uma relação $T(Y)$, onde $Y = Z - X$, que inclui uma tupla t se as tuplas t_R aparecerem em R com $t_R[Y] = t$, e com $t_R[X] = t_S$ para cada tupla t_S em S
 - ▶ Em outras palavras: para uma tupla t aparecer no resultado T da divisão, os valores em t deverão aparecer em R em combinação **com cada tupla** em S .

A operação de Divisão (\div)

Exemplo:

$$T \leftarrow R \div S$$

R

A	B
a1	b1
a2	b1
a3	b1
a4	b1
a1	b2
a3	b2
a2	b3
a3	b3
a4	b3
a1	b4
a2	b4
a3	b4

S

A
a1
a2
a3

T

B
b1
b4

A operação de Divisão (\div)

- ▶ Útil em consultas como a mostrada a seguir:

*Recupere os nomes dos funcionários que trabalham em **todos** os projetos em que 'João Silva' trabalha.*

```
SILVA  $\leftarrow \sigma_{P_{\text{nome}} = \text{'João'} \text{ AND } U_{\text{nome}} = \text{'Silva'}}(\text{FUNCIONARIO})$   
SILVA_PNRS  $\leftarrow \pi_{P_{\text{nr}}}(\text{TRABALHA\_EM} \bowtie_{F_{\text{cpf}} = C_{\text{pf}}} \text{SILVA})$   
CPF_PNRS  $\leftarrow \pi_{F_{\text{cpf}}, P_{\text{nr}}}(\text{TRABALHA\_EM})$   
CPFS( $C_{\text{pf}}$ )  $\leftarrow \text{CPF\_PNRS} \div \text{SILVA\_PNRS}$   
RESULTADO  $\leftarrow \pi_{P_{\text{nome}}, U_{\text{nome}}}(\text{CPFS} * \text{FUNCIONARIO})$ 
```

A operação de Divisão (\div)

- ▶ A divisão pode ser expressa como uma sequência de operações $\pi, \times, -$

$$T \leftarrow R \div S$$

equivale a

$$T1 \leftarrow \pi_Y(R)$$

$$T2 \leftarrow \pi_Y((S \times T1) - R)$$

$$T \leftarrow T1 - T2$$

lembrando que Y é conjunto de atributos que aparece R mas não em S .

Outras operações da Álgebra Relacional

- ▶ Alguns tipos comuns de operações de recuperação de dados em BDs não podem ser realizadas na Álgebra Relacional básica
- ▶ Operações adicionais foram definidas para melhorar o poder expressivo da linguagem
 - ▶ projeção generalizada
 - ▶ funções de agregação e agrupamento
 - ▶ operações de junção externa

Projeção Generalizada

- ▶ Um tipo de projeção que permite que funções sobre os atributos sejam incluídas na lista de projeção
- ▶ Formato:

$$\pi_{F_1, F_2, \dots, F_n}(R)$$

onde

- ▶ F_1, F_2, \dots, F_n são funções sobre os atributos de R
- ▶ As funções podem envolver operações aritméticas e valores constantes

Projeção Generalizada

- ▶ Exemplo:

FUNCIONARIO(Cpf,Salario, Deducao, Anos_em_servico)

RELATORIO $\leftarrow \rho(\text{Cpf, Salario_liquido, Bonus, Imposto})$

$(\pi_{\text{Cpf, Salario - Deducao, } 2000 * \text{Anos_em_servico, } 0.25 * \text{Salario}}$
 $(\text{FUNCIONARIO}))$

Funções de Agregação

- ▶ São funções matemáticas que se aplicam sobre coleções de valores do banco de dados
- ▶ São usadas em consultas estatísticas simples
- ▶ Exemplos
 - ▶ média ou salário total de todos os funcionários
 - ▶ número total de tuplas de projetos
- ▶ Funções comumente aplicadas: SOMA, MÉDIA, MÁXIMO, MÍNIMO, CONTA

Agrupamento

- ▶ Funções de agregação também podem ser aplicadas de forma independente sobre **grupos de tuplas**
- ▶ As tuplas de uma relação podem ser agrupadas segundo o valor de seus atributos
- ▶ Exemplo: as tuplas de FUNCIONARIO podem ser agrupadas por Dnr
 - ▶ cada grupo contém as tuplas dos funcionários que trabalham no mesmo departamento

Operação Função Agregada

- ▶ Combina as operações de agrupamento e agregação
- ▶ Não há uma notação padrão para a operação
- ▶ Notação sugerida:

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

onde

- ▶ $\langle \text{atributos de agrupamento} \rangle$ é uma lista de atributos de R
- ▶ $\langle \text{lista de funções} \rangle$ é uma lista de pares ($\langle \text{função de agregação} \rangle \langle \text{atributo} \rangle$)

Obs.: Em geral, as duplicatas não são eliminadas quando uma função de agregação é aplicada.

Operação Função Agregada

Exemplo 1 (com o resultado renomeado)

FUNCIONARIO

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
João	B	Silva	123456789000	09-01-1965	Rua das Flores, 751, São Paulo, SP	M	30.000	33344555587	5
Fernando	T	Wong	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	M	40.000	88866555576	5
Alice	J	Zelaya	99988777767	19-01-1968	Rua Souza Lima, 35, Curitiba, PR	F	25.000	98765432188	4
Jennifer	S	Souza	98765432188	20-06-1941	Av. Arthur de Lima, 54, Santo André, SP	F	43.000	88866555576	4
Ronaldo	K	Lima	66688444476	15-09-1982	Rua Rebouças, 65, Piracicaba, SP	M	38.000	33344555587	5
Joice	A	Leite	45345345376	31-07-1972	Av. Lucas Obes, 74, São Paulo, SP	F	25.000	33344555587	5
André	V	Pereira	98798798733	29-03-1989	Rua Timbira, 35, São Paulo, SP	M	25.000	98765432188	4
Jorge	E	Brito	88866555576	10-11-1937	Rua do Horto, 35, São Paulo, SP	M	55.000	NULL	1

ρ R(Dnr,Nr_de_funcionarios,Media_sal)(Dnr \bowtie CONTA Cpf, MÉDIA Salario(FUNCIONARIO))

R

Dnr	Nr_de_funcionarios	Media_sal
5	4	33.250
4	3	31.000
1	1	55.000

Resultado

Operação Função Agregada

Exemplo 2

FUNCIONARIO

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
João	B	Silva	12345678966	09-01-1965	Rua das Flores, 751, São Paulo, SP	M	30.000	33344555587	5
Fernando	T	Wong	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	M	40.000	88866555576	5
Alice	J	Zelaya	99988777767	19-01-1968	Rua Souza Lima, 35, Curitiba, PR	F	25.000	98765432168	4
Jennifer	S	Souza	98765432168	20-06-1941	Av. Arthur de Lima, 54, Santo André, SP	F	43.000	88866555576	4
Ronaldo	K	Lima	66688444476	15-09-1962	Rua Rebouças, 65, Piracicaba, SP	M	38.000	33344555587	5
Joice	A	Leite	45345345376	31-07-1972	Av. Lucas Obes, 74, São Paulo, SP	F	25.000	33344555587	5
André	V	Pereira	98798798733	29-03-1969	Rua Timbira, 35, São Paulo, SP	M	25.000	98765432168	4
Jorge	E	Brito	88866555576	10-11-1937	Rua do Horto, 35, São Paulo, SP	M	55.000	NULL	1

$Dnr \overset{S}{\Sigma} \text{CONTA Cpf, MÉDIA Salario}(\text{FUNCIONARIO})$

Dnr	Contador_cpf	Media_salario
5	4	33.250
4	3	31.000
1	1	55.000

Resultado

Operação Função Agregada

Exemplo 3 (sem agrupamento dos dados)

FUNCIONARIO

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
João	B	Silva	12345678966	09-01-1965	Rua das Flores, 751, São Paulo, SP	M	30.000	33344555587	5
Fernando	T	Wong	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	M	40.000	88866555576	5
Alice	J	Zelaya	99988777767	19-01-1968	Rua Souza Lima, 35, Curitiba, PR	F	25.000	98765432168	4
Jennifer	S	Souza	98765432168	20-06-1941	Av. Arthur de Lima, 54, Santo André, SP	F	43.000	88866555576	4
Ronaldo	K	Lima	66688444476	15-09-1962	Rua Rebouças, 65, Piracicaba, SP	M	38.000	33344555587	5
Joice	A	Leite	45345345376	31-07-1972	Av. Lucas Obes, 74, São Paulo, SP	F	25.000	33344555587	5
André	V	Pereira	98798798733	29-03-1969	Rua Timbira, 35, São Paulo, SP	M	25.000	98765432168	4
Jorge	E	Brito	88866555576	10-11-1937	Rua do Horto, 35, São Paulo, SP	M	55.000	NULL	1

$\sigma_{\text{CONTA Cpf, MÉDIA Salario}}(\text{FUNCIONARIO})$

Contador_cpf	Media_salario
8	35.125

Resultado

Operações de Junção Externa

- ▶ Nas operações de junção vistas anteriormente, tuplas sem uma tupla relacionada ou com valores NULL nos atributos de junção são eliminadas do resultado
 - ▶ Essas junções são chamadas de **junções internas** (*inner joins*)
- ▶ As **junções externas** permitem manter todas as tuplas “espúrias” que se desejar no conjunto resposta. Por exemplo, em um junção externa entre as relações R e S é possível manter no resultado:
 - ▶ todas as tuplas de R (junção externa à esquerda), ou
 - ▶ todas as tuplas de S (junção externa à direita), ou
 - ▶ todas as tuplas de R e de S (junção externa completa)independentemente de elas possuírem ou não tuplas associadas na outra relação.

Operação de Junção Externa à Esquerda

- ▶ Notação:

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

- ▶ A operação mantém cada tupla da relação à esquerda (R)
- ▶ Se nenhuma tupla correspondente for encontrada em S , então os atributos de S no resultado da junção são preenchidos com valores NULL

De forma análoga funcionam as operações de:

- ▶ Junção Externa à Direita: $R \bowtie_{\langle \text{condição de junção} \rangle} S$
(que mantém todas as tuplas de S)
- ▶ Junção Externa Completa: $R \bowtie_{\langle \text{condição de junção} \rangle} S$
(que mantém todas as tuplas de R e de S)

Operação de Junção Externa à Esquerda

Exemplo: $R \bowtie_{B=D \text{ AND } C=E} S$

A	B	C
1	2	3
4	5	6
7	8	9

Relação R

D	E	F
2	3	10
2	3	11
6	7	12

Relação S

A	B	C	D	E	F
1	2	3	2	3	10
1	2	3	2	3	11
4	5	6	NULL	NULL	NULL
7	8	9	NULL	NULL	NULL

Resultado de $R \bowtie_{B=D \text{ AND } C=E} S$

Operação de Junção Externa à Direita

Exemplo: $R \bowtie_{B=D \text{ AND } C=E} S$

A	B	C
1	2	3
4	5	6
7	8	9

Relação R

D	E	F
2	3	10
2	3	11
6	7	12

Relação S

A	B	C	D	E	F
1	2	3	2	3	10
1	2	3	2	3	11
NULL	NULL	NULL	6	7	12

Resultado de $R \bowtie_{B=D \text{ AND } C=E} S$

Operação de Junção Externa Completa

Exemplo: $R \bowtie_{B=D \text{ AND } C=E} S$

A	B	C
1	2	3
4	5	6
7	8	9

Relação R

D	E	F
2	3	10
2	3	11
6	7	12

Relação S

A	B	C	D	E	F
1	2	3	2	3	10
1	2	3	2	3	11
4	5	6	NULL	NULL	NULL
7	8	9	NULL	NULL	NULL
NULL	NULL	NULL	6	7	12

Resultado de $R \bowtie_{B=D \text{ AND } C=E} S$

Operação de Junção Externa à Esquerda

Exemplo

$$\text{TEMP} \leftarrow (\text{FUNCIONARIO} \bowtie_{\text{Cpf} = \text{Cpf_gerente}} \text{DEPARTAMENTO})$$

$$\text{RESULTADO} \leftarrow \pi_{\text{Pnome}, \text{Minicial}, \text{Unome}, \text{Dnome}}(\text{TEMP})$$

RESULTADO

Pnome	Minicial	Unome	Dnome
João	B	Silva	NULL
Fernando	T	Wong	Pesquisa
Alice	J	Zelaya	NULL
Jennifer	S	Souza	Administração
Ronaldo	K	Lima	NULL
Joice	A	Leite	NULL
André	V	Pereira	NULL
Jorge	E	Brito	Matriz

Referências Bibliográficas

- ▶ *Sistemas de Bancos de Dados* (6ª edição), Elmasri e Navathe. Pearson, 2010. – Capítulo 6
- ▶ *Database Systems – The Complete Book*, Garcia-Molina, Ulmann e Widom. Prentice Hall, 2002. – Capítulo 5

Cenas dos próximos capítulos...

- ▶ Cálculo de tuplas e de domínio