

[MAC0426] Sistemas de Bancos de Dados
[IBI5013] Bancos de Dados para Bioinformática
Aula 10
A Álgebra Relacional (Parte 1)

Kelly Rosa Braghetto

DCC-IME-USP

20 de abril de 2017

Sobre a Álgebra Relacional e o Cálculo Relacional

- ▶ São linguagens formais para o modelo relacional
- ▶ São a base da linguagem SQL
- ▶ Permitem a definição de expressões de recuperação de dados em relações

A Álgebra Relacional

- ▶ É considerada uma parte integral do modelo de dados relacional
- ▶ É usada como base para a implementação e otimização de consultas nos SGBDRs
- ▶ Suas operações podem ser divididas em dois grupos:
 - ▶ operações da teoria de conjuntos (matemática) – **união, intersecção, diferença de conjunto e produto cartesiano**
 - ▶ operações específicas para BD relacionais – **seleção, projeção, junção, divisão**

Operações unárias

Seleção

- ▶ Operação usada para escolher um **subconjunto** das tuplas (linhas) de uma relação
- ▶ As tuplas selecionadas são as que satisfazem a **condição de seleção**
- ▶ Pode ser vista como uma partição horizontal da relação em dois conjuntos:
 - ▶ as tuplas que satisfazem a condição e são selecionadas
 - ▶ as tuplas que não satisfazem a condição e são descartadas

Operações unárias

Seleção

É expressa como

$$\sigma_{\langle \text{condição de seleção} \rangle}(R)$$

onde

- ▶ o símbolo σ indica o operador de seleção
- ▶ R é o resultado de uma **expressão da álgebra relacional** (que é sempre uma relação!)
 - ▶ A expressão mais simples desse tipo é apenas o nome de uma relação do banco de dados
- ▶ a condição de seleção é uma expressão booleana especificada sobre os atributos de R

Operações unárias

Seleção – Exemplos

- ▶ Seleção das tuplas de FUNCIONARIO cujo departamento é 4:

$$\sigma_{Dnr = 4}(\text{FUNCIONARIO})$$

- ▶ Seleção das tuplas de FUNCIONARIO cujo salário é maior do que R\$ 30.000,00:

$$\sigma_{\text{Salario} > 30.000}(\text{FUNCIONARIO})$$

Operações unárias

Seleção

- ▶ A condição de seleção é uma expressão booleana composta por **cláusulas** no formato

<atributo> <op. de comparação> <valor constante>

ou

<atributo> <op. de comparação> <atributo>

- ▶ **<atributo>** – nome de um atributo de R
- ▶ **<op. de comparação>** – $=, <, \leq, >, \geq$ ou \neq
- ▶ **<valor constante>** – um valor constante do domínio do atributo ao qual está sendo comparado

Obs.: Os operadores $<, \leq, >, \geq$ só podem ser usados com atributos cujo domínio são conjuntos de **valores ordenados** (como números, datas, cadeias de caracteres, etc.). Para domínios de valores desordenados, apenas os operadores de comparação de igualdade ($=, \neq$) podem ser usados.

Operações unárias

Seleção – condição

- ▶ Cláusulas podem ser conectadas para formar uma condição por meio de operadores lógicos:
 - ▶ **and, or, not**
- ▶ A condição de seleção não pode envolver mais de uma tupla → a operação de seleção é aplicada a cada tupla individualmente
- ▶ Ex: Seleção das tuplas de funcionários que ou trabalham no departamento 4 e ganham mais do que R\$25.000,00, ou trabalham no departamento 5 e ganham mais do que R\$ 30.000,00

$\sigma_{(Dnr = 4 \text{ AND } Salario > 25.000) \text{ OR } (Dnr = 5 \text{ AND } Salario > 30.000)}(\text{FUNCIONARIO})$

Relembrando o esquema do BD Empresa

FUNCIONARIO

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

DEPARTAMENTO

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

LOCALIZACAO_DEP

<u>Dnumero</u>	<u>Dlocal</u>
----------------	---------------

PROJETO

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

TRABALHA_EM

<u>Fopf</u>	<u>Pnr</u>	Horas
-------------	------------	-------

DEPENDENTE

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

Operações unárias

Seleção – condição

FUNCIONARIO

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
João	B	Silva	123456789066	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

$\sigma(\text{Dnr} = 4 \text{ AND Salario} > 25.000) \text{ OR } (\text{Dnr} = 5 \text{ AND Salario} > 30.000)(\text{FUNCIONARIO})$

Resultado

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
Fernando	T	Wong	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	M	40.000	88866555576	5
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

Operações unárias

Seleção – propriedades

- ▶ O operador de seleção é comutativo:

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

- ▶ Uma sequência de operações de seleção pode ser substituída por uma única operação de seleção com condição conjuntiva (AND):

$$\begin{aligned} \sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(\dots(\sigma_{\langle \text{condn} \rangle}(R))\dots)) = \\ \sigma_{\langle \text{cond1} \rangle \text{ AND } \langle \text{cond2} \rangle \text{ AND } \dots \text{ AND } \langle \text{condn} \rangle}(R) \end{aligned}$$

Operações unárias

Seleção – exemplo

- ▶ Seleção das funcionárias que trabalham no departamento 4 e ganham mais do que R\$25.000,00

$$\sigma_{\text{Salario} > 25.000}(\sigma_{\text{Dnr} = 4}(\sigma_{\text{Sexo} = 'F'}(\text{FUNCIONARIO})))$$

ou

$$\sigma_{\text{Sexo} = 'F' \text{ AND } \text{Dnr} = 4 \text{ AND } \text{Salario} > 25.000}(\text{FUNCIONARIO})$$

Operações unárias

Projeção

- ▶ Operação que seleciona atributos (colunas) de uma relação, projetando a relação sobre esses atributos
- ▶ Pode ser vista como uma partição vertical da relação em duas relações:
 - ▶ uma relação tem os atributos necessários e contém o resultado da operação
 - ▶ a outra relação contém os atributos descartados

Operações unárias

Projeção

- ▶ É expressa como

$$\pi_{\langle \text{lista de atributos} \rangle}(R)$$

onde

- ▶ o símbolo π expressa a operação de projeção
 - ▶ R é uma **expressão da álgebra relacional**
-
- ▶ Resultado de uma projeção: uma relação que contém apenas os atributos especificados na $\langle \text{lista de atributos} \rangle$, **exatamente na mesma ordem em que eles aparecem na lista**
 - ▶ Ex: Listar último nome, primeiro nome e salário de cada funcionário

$$\pi_{\text{Unome, Pnome, Salario}}(\text{FUNCIONARIO})$$

Operações unária – Projeção

FUNCIONARIO

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
João	B	Silva	123456789066	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

$\pi_{\text{Unome, Pnome, Salario}}(\text{FUNCIONARIO})$

Unome	Pnome	Salario
Silva	João	30.000
Wong	Fernando	40.000
Zelaya	Alice	25.000
Souza	Jennifer	43.000
Lima	Ronaldo	38.000
Leite	Joice	25.000
Pereira	André	25.000
Brito	Jorge	55.000

Resultado:

Operações unárias

Projeção

- ▶ Se a lista de atributos da projeção não contém atributos chave de R , a projeção pode resultar tuplas duplicadas
- ▶ A operação de projeção **sempre remove as tuplas duplicadas do conjunto de resposta** → o resultado da operação é sempre uma relação válida

Operações unárias

Projeção – exemplo de remoção de tuplas duplicadas

FUNCIONARIO

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
João	B	Silva	123456789066	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

$\pi_{\text{Sexo, Salario}}(\text{FUNCIONARIO})$

Sexo	Salario
M	30.000
M	40.000
F	25.000
F	43.000
M	38.000
M	25.000
M	55.000

Resultado:

Operações unárias

Projeção – propriedades

- ▶ A comutatividade não vale para a operação de projeção
- ▶ A propriedade a seguir vale quando $\langle \text{lista1} \rangle \subseteq \langle \text{lista2} \rangle$

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

- ▶ Se $\langle \text{lista1} \rangle \not\subseteq \langle \text{lista2} \rangle$, então a expressão abaixo é incorreta

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

Sequências de operações

Ex.: Recuperar o primeiro nome, o último nome e o salário de todos os funcionários que trabalham no departamento 5.

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

$$\pi_{\text{Pnome, Unome, Salario}}(\sigma_{\text{Dnr}=5}(\text{FUNCIONARIO}))$$

Pnome	Unome	Salario
João	Silva	30.000
Fernando	Wong	40.000
Ronaldo	Lima	38.000
Joice	Leite	25.000

Resultado:

Sequências de operações e a operação Renomear (\leftarrow)

- ▶ Às vezes, é conveniente desmembrar uma expressão complexa em expressões mais simples, que geram relações com resultados intermediários às quais podemos atribuir nomes.
- ▶ Exemplo:

$$\pi_{\text{Pnome, Unome, Salario}}(\sigma_{\text{Dnr}=5}(\text{FUNCIONARIO}))$$

equivale a

$$\text{FUNCS_DEPT5} \leftarrow \sigma_{\text{Dnr}=5}(\text{FUNCIONARIO})$$

$$\text{RESULTADO} \leftarrow \pi_{\text{Pnome, Unome, Salario}}(\text{FUNCS_DEPT5})$$

A operação renomear (\leftarrow)

- ▶ A operação de renomear pode ser usada também para renomear os atributos
- ▶ Ex. (os novos nomes de atributos aparecem entre parênteses):

$$\text{TEMP} \leftarrow \sigma_{\text{Dnr}=5}(\text{FUNCIONARIO})$$

$$R(\text{Prim_nome}, \text{Ult_nome}, \text{Salario}) \leftarrow \pi_{\text{Pnome}, \text{Unome}, \text{Salario}}(\text{TEMP})$$

- ▶ Existe também uma formalização da operação renomear como o operador unário ρ , que permite mudar o nome da relação, ou os nomes dos atributos, ou as duas coisas ao mesmo tempo:

$$\rho_S(R) \text{ ou } \rho_{(B_1, B_2, \dots, B_n)}(R) \text{ ou } \rho_{S(B_1, B_2, \dots, B_n)}(R)$$

A operação renomear (\leftarrow)

$$\text{TEMP} \leftarrow \sigma_{\text{Dnr}=5}(\text{FUNCIONARIO})$$

$$\text{R}(\text{Primeiro_nome}, \text{Ultimo_nome}, \text{Salario}) \leftarrow$$

$$\pi_{\text{Pnome}, \text{Unome}, \text{Salario}}(\text{TEMP})$$

Resultado:

TEMP

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
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

R

Primeiro_nome	Ultimo_nome	Salario
João	Silva	30.000
Fernando	Wong	40.000
Ronaldo	Lima	38.000
Joice	Leite	25.000

Operações da teoria dos conjuntos

União

- ▶ A expressão $R \cup S$ denota a relação formada pelas tuplas que estão ou em R , ou em S , ou em ambas
 - ▶ A operação de união **elimina as tuplas duplicadas** \rightarrow o resultado é sempre uma relação válida

Intersecção

- ▶ A expressão $R \cap S$ denota a relação formada pelas tuplas que estão tanto em R quanto em S

Diferença (ou Subtração)

- ▶ A expressão $R - S$ denota a relação formada pelas tuplas que estão em R mas não em S

Operações da teoria dos conjuntos

Exemplo da União

ALUNO

Pn	Un
Susana	Yao
Ronaldo	Lima
José	Gonçalves
Barbara	Pires
Ana	Tavares
Jonas	Wang
Ernesto	Gilberto

PROFESSOR

Pnome	Unome
João	Silva
Ricardo	Braga
Susana	Yao
Franiscoo	Leme
Ronaldo	Lima

Pn	Un
Susana	Yao
Ronaldo	Lima
José	Gonçalves
Barbara	Pires
Ana	Tavares
Jonas	Wang
Ernesto	Gilberto
João	Silva
Ricardo	Braga
Franiscoo	Leme

ALUNO \cup PROFESSOR

Operações da teoria dos conjuntos

Exemplo da Intersecção

ALUNO

Pn	Un
Susana	Yao
Ronaldo	Lima
José	Gonçalves
Barbara	Pires
Ana	Tavares
Jonas	Wang
Ernesto	Gilberto

PROFESSOR

Pnome	Unome
João	Silva
Ricardo	Braga
Susana	Yao
Francoisco	Leme
Ronaldo	Lima

ALUNO \cap PROFESSOR

Pn	Un
Susana	Yao
Ronaldo	Lima

Operações da teoria dos conjuntos

Exemplo da Subtração

ALUNO

Pn	Un
Susana	Yao
Ronaldo	Lima
José	Gonçalves
Barbara	Pires
Ana	Tavares
Jonas	Wang
Ernesto	Gilberto

PROFESSOR

Pnome	Unome
João	Silva
Ricardo	Braga
Susana	Yao
Franiscoo	Leme
Ronaldo	Lima

Pn	Un
José	Gonçalves
Barbara	Pires
Ana	Tavares
Jonas	Wang
Ernesto	Gilberto

ALUNO – PROFESSOR

Operações da teoria dos conjuntos

Exemplo da Subtração

ALUNO

Pn	Un
Susana	Yao
Ronaldo	Lima
José	Gonçalves
Barbara	Pires
Ana	Tavares
Jonas	Wang
Ernesto	Gilberto

PROFESSOR

Pnome	Unome
João	Silva
Ricardo	Braga
Susana	Yao
Franiscoo	Leme
Ronaldo	Lima

PROFESSOR – ALUNO

Pnome	Unome
João	Silva
Ricardo	Braga
Franiscoo	Leme

Operações da teoria dos conjuntos

União, intersecção e diferença

- ▶ São operações binárias
- ▶ Só se aplicam a pares de relações que possuem **o mesmo tipo de tuplas** (propriedade chamada de *compatibilidade de união* ou *compatibilidade de tipo*)
- ▶ Duas relações $R(A_1, A_2, \dots, A_n)$ e $S(B_1, B_2, \dots, B_m)$ são compatíveis se $n = m$ e se $dom(A_i) = dom(B_i)$ para $1 \leq i \leq n$
- ▶ Convenção: a relação resultante tem os mesmos nomes de atributos da primeira relação R

Operações da teoria dos conjuntos

União e intersecção

- ▶ São comutativas

$$R \cup S = S \cup R$$

$$R \cap S = S \cap R$$

- ▶ São associativas

$$R \cup (S \cup T) = (S \cup R) \cup T$$

$$R \cap (S \cap T) = (S \cap R) \cap T$$

Subtração

- ▶ Não é comutativa; ou seja, em geral $R - S \neq S - R$

$$R \cap S = ((R \cup S) - (R - S)) - (S - R) = R - (R - S)$$

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>:

$\langle \text{condição1} \rangle \text{ AND } \langle \text{condição2} \rangle \text{ AND } \dots \text{ AND } \langle \text{condição}p \rangle$

onde cada $\langle \text{condição}k \rangle$, $1 \leq k \leq p$, tem a forma $A_i \text{ op } B_j$ 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: $R \bowtie_{\langle \text{condição} \rangle} \equiv \sigma_{\langle \text{condição} \rangle}(R \times S)$

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  $\leftarrow \sigma_{Pnome = 'João' \text{ AND } Unome = 'Silva'}(\text{FUNCIONARIO})$   
SILVA_PNRS  $\leftarrow \pi_{Pnr}(\text{TRABALHA\_EM} \bowtie_{Fcpf = Cpf} \text{SILVA})$   
CPF_PNRS  $\leftarrow \pi_{Fcpf, Pnr}(\text{TRABALHA\_EM})$   
CPFS(Cpf)  $\leftarrow \text{CPF\_PNRS} \div \text{SILVA\_PNRS}$   
RESULTADO  $\leftarrow \pi_{Pnome, Unome}(\text{CPFS} * \text{FUNCIONARIO})$ 
```

A operação de Divisão (\div)

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

$$T \leftarrow R \div S$$

equivale a

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

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

$$T \leftarrow T1 - T2$$

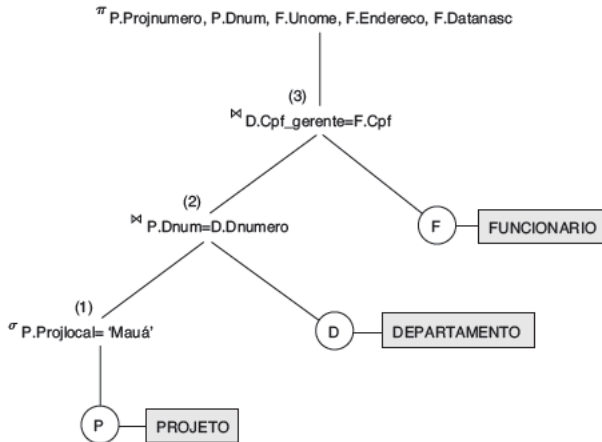
lembrando que Y é conjunto de atributos que aparece R mas não em S .

Árvore de consulta

- ▶ É uma estrutura de dados em árvore que corresponde a uma expressão da álgebra relacional
 - ▶ **nós folhas** – relações de entrada da consulta
 - ▶ **nós internos** – operações da álgebra relacional
- ▶ Com frequência é usada por SGBDRs para a representação interna de consultas
- ▶ Oferece uma boa representação visual e compreensão da consulta

Árvore de consulta

Exemplo: Para cada projeto localizado em 'Mauá', liste o número do projeto, o número do departamento que o controla, e o último nome, endereço e data de nascimento do gerente do departamento.



Ferramentas para executar consultas em álgebra relacional

- ▶ **RelaX** – *relational algebra calculator* (ferramenta online, criada na Universidade de Innsbruck)
<http://dbis-uibk.github.io/relax/>
- ▶ **WinRDBI Educational Tool** (ferramenta em Java, criada na Arizona State University)
<https://winrdbi.asu.edu/>
- ▶ **Relational** – *Educational tool for relational algebra*
<http://ltworf.github.io/relational/>

Atenção: essas ferramentas **não** são SGBDRs. Elas são somente ferramentas para o ensino/aprendizagem de Álgebra Relacional, possibilitando que consultas sejam executadas sobre pequenos conjuntos de dados relacionais.

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...

Mais sobre linguagens de consulta

- ▶ Álgebra Relacional (Parte 2)
- ▶ Cálculo de tuplas