

[MAC0426] Sistemas de Bancos de Dados
[IBI5013] Bancos de Dados para Bioinformática
Aula 10
A Álgebra Relacional (Parte 2)

Kelly Rosa Braghetto

DCC-IME-USP

1º de abril de 2016

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

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

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

$$T \leftarrow T1 - T2$$

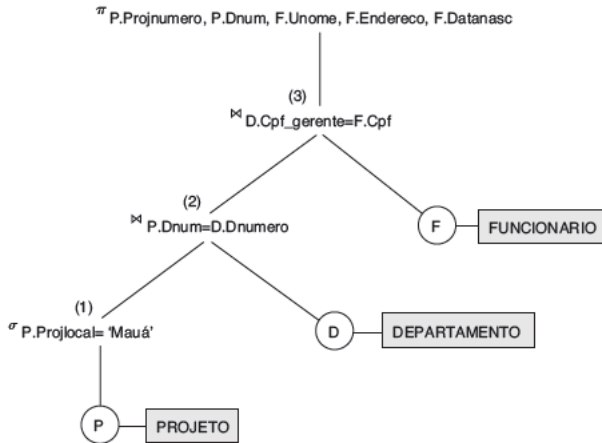
lembrando que Y é conjunto de atributos que aparece R mas não em S .

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.



Outras operações da Álgebra Relacional

- ▶ Alguns tipos comuns de operações de recuperação de dados em BDs não podem ser realizados 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)

- ▶ Queremos mostrar:

Salário líquido \rightarrow Salário - Dedução

Bônus $\rightarrow 2.000,00 * \text{Anos em serviço}$

Imposto $\rightarrow 0,25 * \text{Salário}$

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:

$$\langle \text{atributos de agrupamento} \rangle \mathfrak{S} \langle \text{lista de funções} \rangle (R)$$

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	29-03-1969	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	12345678900	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	66688444478	15-09-1962	Rua Rebouças, 65, Piracicaba, SP	M	38.000	33344555587	5
Joice	A	Leite	45345345378	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 desejam 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

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

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

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

Essa consulta devolve o nome completo de todos os funcionários e, para os funcionários que gerenciam um departamento, devolve também o nome do departamento que gerenciam.

Exemplos de consultas em Álgebra Relacional

FUNCIONARIO

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

DEPARTAMENTO

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

LOCALIZACOES_DEP

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

PROJETO

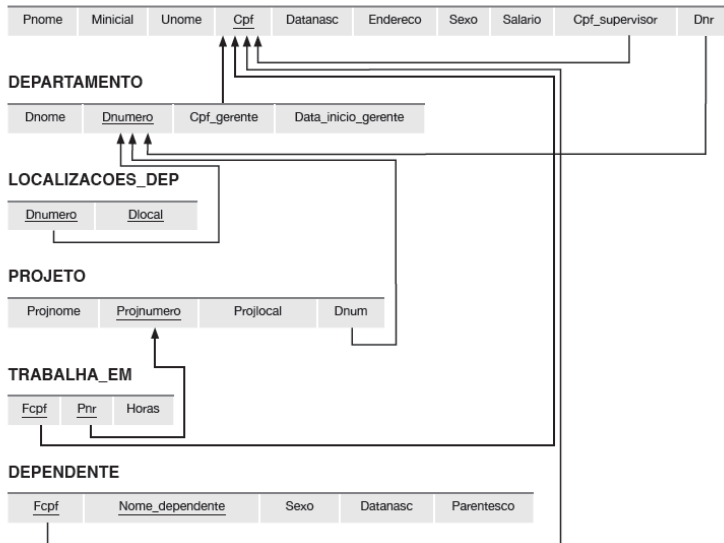
Projnome	<u>Projnumero</u>	Projlocal	Dnum
----------	-------------------	-----------	------

TRABALHA_EM

<u>Fcpf</u>	<u>Pnr</u>	Horas
-------------	------------	-------

DEPENDENTE

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



Exemplos de consultas em Álgebra Relacional

Consulta 1: Recuperar o nome e o endereço de todos os funcionários que trabalham para o departamento 'Pesquisa'.

Exemplos de consultas em Álgebra Relacional

Consulta 1: Recuperar o nome e o endereço de todos os funcionários que trabalham para o departamento 'Pesquisa'.

$DEP_PESQUISA \leftarrow \sigma_{Dnome = 'Pesquisa'}(DEPARTAMENTO)$

$FUNCS_PESQUISA \leftarrow DEP_PESQUISA \bowtie_{Dnumero = Dnr} FUNCIONARIO$

$RESULTADO \leftarrow \pi_{Pnome, Unome, Endereco}(FUNCS_PESQUISA)$

ou, numa só expressão:

$\pi_{Pnome, Unome, Endereco}((\sigma_{Dnome = 'Pesquisa'}(DEPARTAMENTO)) \bowtie_{Dnumero = Dnr} FUNCIONARIO)$

Exemplos de consultas em Álgebra Relacional

Consulta 2: Descobrir o nome dos funcionários que trabalham em *todos* os projetos controlados pelo departamento número 5.

Exemplos de consultas em Álgebra Relacional

Consulta 2: Descobrir o nome dos funcionários que trabalham em *todos* os projetos controlados pelo departamento número 5.

$$\text{PROJ_DEP5} \leftarrow \rho_{\text{Pnr}}(\pi_{\text{Projnumero}}(\sigma_{\text{Dnum} = 5}(\text{PROJETO})))$$
$$\text{FUNC_PROJ} \leftarrow \rho_{\text{Cpf, Pnr}}(\pi_{\text{Fcpf, Pnr}}(\text{TRABALHA_EM}))$$
$$\text{RESULTADO_CPF_FUNC} \leftarrow \text{FUNC_PROJ} \div \text{PROJ_DEP5}$$
$$\text{RESULTADO} \leftarrow$$
$$\pi_{\text{Unome, Pnome}}(\text{RESULTADO_CPF_FUNC} * \text{FUNCIONARIO})$$

Exemplos de consultas em Álgebra Relacional

Consulta 3: Fazer uma lista dos números de projeto para aqueles que envolvem um funcionário cujo último nome é 'Silva', seja como um trabalhador ou como um gerente do departamento que controla o projeto.

Exemplos de consultas em Álgebra Relacional

Consulta 3: Fazer uma lista dos números de projeto para aqueles que envolvem um funcionário cujo último nome é 'Silva', seja como um trabalhador ou como um gerente do departamento que controla o projeto.

$$\text{SILVA}(\text{Fcpf}) \leftarrow (\pi_{\text{Ccpf}}(\sigma_{\text{Unome} = \text{'Silva'}}(\text{FUNCIONARIO})))$$

$$\text{PROJS_SILVA_TRABALHA} \leftarrow \pi_{\text{Pnr}}(\text{TRABALHA_EM} * \text{SILVA})$$

$$\text{DEPS_GERENCIADOS_SILVA}(\text{Dnum}) \leftarrow$$

$$\pi_{\text{Dnumero}}(\text{SILVA} \bowtie_{\text{Fcpf}=\text{Cpf_gerente}} \text{DEPARTAMENTO})$$

$$\text{PROJS_SILVA_GER}(\text{Pnr}) \leftarrow$$

$$\pi_{\text{Projnumero}}(\text{DEPS_GERENCIADOS_SILVA} * \text{PROJETO})$$

$$\text{RESULTADO} \leftarrow$$

$$\text{PROJS_SILVA_TRABALHA} \cup \text{PROJS_SILVA_GER}$$

Exemplos de consultas em Álgebra Relacional

Consulta 4: Listar o nome dos funcionários com dois ou mais dependentes.

Exemplos de consultas em Álgebra Relacional

Consulta 4: Listar o nome dos funcionários com dois ou mais dependentes.

$$T1(\text{Cpf}, \text{Nr_dependentes}) \leftarrow_{F_{\text{cpf}}} \mathcal{S}_{\text{CONTA Nome_dependente}}(\text{DEPENDENTE})$$

$$T2 \leftarrow \sigma_{\text{Nr_dependentes} > 2}(T1)$$

$$\text{RESULTADO} \leftarrow \pi_{\text{Unome}, \text{Pnome}}(T2 * \text{FUNCIONARIO})$$

ou (sem usar função agregada):

$$T1(\text{Fcpf1}, \text{Dependente1}) \leftarrow \pi_{F_{\text{cpf}}, \text{Nome_dependente}}(\text{DEPENDENTE})$$

$$T2(\text{Fcpf2}, \text{Dependente2}) \leftarrow \pi_{F_{\text{cpf}}, \text{Nome_dependente}}(\text{DEPENDENTE})$$

$$T3(\text{Cpf}) \leftarrow \pi_{F_{\text{cpf}}}(T1 \bowtie_{F_{\text{cpf1}} = F_{\text{cpf2}} \text{ and } \text{Dependente1} \neq \text{Dependente2}} T2)$$

$$\text{RESULTADO} \leftarrow \pi_{\text{Unome}, \text{Pnome}} T3 * \text{FUNCIONARIO}$$

Exemplos de consultas em Álgebra Relacional

Consulta 5: Recuperar os nomes dos funcionários que não possuem dependentes.

Exemplos de consultas em Álgebra Relacional

Consulta 5: Recuperar os nomes dos funcionários que não possuem dependentes.

$TODOS_FUNCS \leftarrow \pi_{Cpf}(FUNCIONARIO)$

$FUNCS_COM_DEPEND(Cpf) \leftarrow \pi_{Fcpf}(DEPENDENTE)$

$FUNCS_SEM_DEPEND \leftarrow$

$TODOS_FUNC - FUNCS_COM_DEPEND$

$RESULTADO \leftarrow$

$\pi_{Unome, Pnome} FUNCS_SEM_DEPEND * FUNCIONARIO$

Exemplos de consultas em Álgebra Relacional

Consulta 6: Listar os nomes dos gerentes que têm pelo menos um dependente.

Exemplos de consultas em Álgebra Relacional

Consulta 6: Listar os nomes dos gerentes que têm pelo menos um dependente.

$GERES(Cpf) \leftarrow \pi_{Cpf_gerente}(DEPARTAMENTO)$

$FUNCS_COM_DEPEND(Cpf) \leftarrow \pi_{Fcpf}(DEPENDENTE)$

$GERS_COM_DEPEND \leftarrow GERES \cap FUNCS_COM_DEPEND$

RESULTADO \leftarrow

$\pi_{Unome, Pnome} GERS_COM_DEPEND * FUNCIONARIO$

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

- ▶ Cálculo de tuplas e de domínio