

# Álgebra Relacional

# Modelo Relacional - Manipulação

- Duas categorias de linguagens
  - formais
    - álgebra relacional e cálculo relacional
  - comerciais (baseadas nas linguagens formais)
    - SQL
- Linguagens formais - Características
  - orientadas a conjuntos
  - linguagens de base
    - linguagens relacionais devem ter no mínimo um poder de expressão equivalente ao de uma linguagem formal
  - fechamento
    - resultados de consultas são relações

# Álgebra Relacional

- Operadores para consulta e alteração de relações
- Linguagem procedural
  - uma expressão na álgebra define uma execução seqüencial de operadores
  - a execução de cada operador produz uma relação
- Classificação dos operadores
  - fundamentais
    - unários: *seleção*, *projeção*
    - binários: *produto cartesiano*, *união* e *diferença*
  - derivados
    - binários: *intersecção*, *junção* e *divisão*
  - especiais
    - *renomeação* (unário) e *atribuição*
  - operador de *alteração* (unário)

# Esquema Relacional Exemplo

Ambulatórios(*#nroa*, andar, capacidade)

Médicos(*#codm*, CPF, nome, idade, cidade, especialidade, *&nroa*)

Pacientes(*#codp*, CPF, nome, idade, cidade, doença)

Consultas(*#&codm*, *#&codp*, *#data*, hora)

Funcionários(*#codf*, CPF, nome, idade, cidade, salário)

# Seleção

- Retorna **tuplas** que satisfazem um **predicado**
- Resultado
  - subconjunto horizontal de uma relação
- Notação

$$\sigma_{\text{predicado}}(\text{relação})$$

- Operadores de comparação

$=, <, <=, >, >=, \neq$

- Operadores lógicos:  $\wedge$  (and)  $\vee$  (or)  $\neg$  (not)

- Exemplo:  $\sigma_{z \geq 2}(R)$

*R*

<i>x</i>	<i>y</i>	<i>z</i>
1	1	1
2	2	2
2	2	3

*resultado*

<i>x</i>	<i>y</i>	<i>z</i>
2	2	2
2	2	3

# Para resolver

- 1) buscar os dados dos pacientes que estão com sarampo
- 2) buscar os dados dos médicos ortopedistas com mais de 55 anos
- 3) buscar os dados de todas as consultas, exceto aquelas marcadas para os médicos com código 46 e 79
- 4) buscar os dados dos ambulatorios do quarto andar. Estes ambulatorios devem ter capacidade igual a 50 ou número superior a 10

# Projeção

- Retorna um ou mais **atributos** de interesse
- Resultado
  - subconjunto vertical de uma relação
- Notação

$\pi_{\text{lista\_nomes\_atributos}}(\text{relação})$

- Eliminação automática de duplicatas
- Exemplo:  $\pi_{x,y}(R)$

*R*

<i>x</i>	<i>y</i>	<i>z</i>
1	1	1
2	2	2
2	2	3

*resultado*

<i>x</i>	<i>y</i>
1	1
2	2

# Para resolver

- 1) buscar o nome e a especialidade de todos os médicos
- 2) buscar o número dos ambulatórios do terceiro andar
- 3) buscar o código dos médicos e as datas das consultas para os pacientes com código 122 e 725
- 4) buscar os números dos ambulatórios com capacidade superior a 50, exceto aqueles do segundo e quarto andares



# Produto Cartesiano

- Retorna todas as combinações de tuplas de duas relações  $R_1$  e  $R_2$
- Grau do resultado
  - $\text{grau}(R_1) + \text{grau}(R_2)$
- Cardinalidade do resultado
  - $\text{cardinalidade}(R_1) * \text{cardinalidade}(R_2)$
- Notação

relação1 X relação2

- Exemplo:

$R_1$

$x$	$y$	$z$
1	1	1
2	2	2
3	3	3

$R_2$

$w$	$y$
1	1
2	2

$R_1 \times R_2$

$x$	$R_1.y$	$z$	$w$	$R_2.y$
1	1	1	1	1
1	1	1	2	2
2	2	2	1	1
2	2	2	2	2
3	3	3	1	1
3	3	3	2	2

# Para resolver

- 1) buscar o **nome** dos médicos que têm consulta marcada e as **datas** das suas consultas
- 2) buscar o **número** e a **capacidade** dos ambulatórios do quinto andar e o **nome** dos médicos ortopedistas que atendem neles
- 3) buscar, para as consultas marcadas para o período da manhã (7hs-12hs) do dia 25/09/06, o **nome** do médico, o **nome** do paciente e a **data** da consulta
- 4) buscar o **nome** e o **salário** dos funcionários de Florianópolis e Palhoça que estão internados como pacientes e têm consulta marcada com psiquiatras

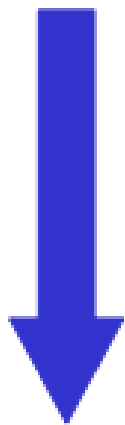
# Otimização Algébrica

- Antecipação de seleções
  - filtragens horizontais o mais cedo possível
- Definição de projeções
  - filtragens verticais o mais cedo possível
    - desde que não prejudiquem operações algébricas futuras que necessitem de atributos eliminados
- Identificação de sub-expressões comuns
  - processá-la uma única vez, mantendo-a em uma variável de relação
  - esta variável de relação é usada várias vezes no processamento da consulta

# Exemplo de Otimização

- Buscar o nome dos médicos que estão internados como pacientes, sofrendo de hepatite

$\pi_{\text{Médicos.nome}} (\sigma_{\text{Pacientes.CPF} = \text{Médicos.CPF} \wedge \text{doença} = \text{'hepatite'}} (\text{Pacientes X Médicos}))$



*antecipando seleções e definindo projeções*

$\pi_{\text{nome}} (\sigma_{\text{Pacientes.CPF} = \text{Médicos.CPF}} (\pi_{\text{CPF}} (\sigma_{\text{doença} = \text{'hepatite'}} (\text{Pacientes}))) \times (\pi_{\text{CPF, nome}} (\text{Médicos})))$

# Renomeação

- Altera o nome de uma relação e/ou dos seus atributos
- Notação

$\rho_{(\text{nome\_atributo1}, \dots, \text{nome\_atributoN})\text{E/OU nome\_relação}}(\text{relação})$

- Exemplos

$R$			$R \times \rho_{R1}(R)$					
$x$	$y$	$z$	$R.x$	$R.y$	$R.z$	$R1.x$	$R1.y$	$R1.z$
1	1	1	1	1	1	1	1	1
2	1	3	1	1	1	2	1	3
			2	1	3	1	1	1
			2	1	3	2	1	3

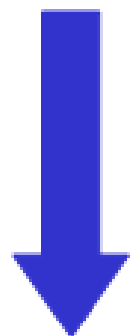
$\rho_{(a, b, c)}(R)$		
$a$	$b$	$c$
1	1	1
2	1	3

# Para resolver

- 1) buscar o **número** dos ambulatórios com capacidade superior à capacidade do ambulatório de número 100
- 2) buscar o **nome** e o **CPF** dos funcionários que recebem salários iguais ou inferiores ao salário do funcionário com CPF 1001
- 3) buscar pares de **nomes** de médicos diferentes que têm consultas marcadas nas mesmas datas

# Exemplo de Otimização

- Buscar o número dos ambulatórios onde pelo menos dois médicos de Florianópolis dão atendimento

$$\pi_{M.nroa} (\sigma_{Médicos.nroa = M.nroa \wedge Médicos.codm \neq M.codm} ((\sigma_{cidade = 'Fpolis'} (Médicos)) \times (\rho_M (\sigma_{cidade = 'Fpolis'} (Médicos)))))$$


*definindo projeções e identificando sub-expressões em comum*

$$R1 \leftarrow \pi_{codm, nroa} (\sigma_{cidade = 'Fpolis'} (Médicos))$$

$$Resposta \leftarrow \pi_{R1.nroa} (\sigma_{R1.nroa = R2.nroa \wedge R1.codm \neq R2.codm} (R1 \times \rho_{R2}(R1)))$$

# União, Diferença e Intersecção

- Operam **somente** sobre duas relações  $R_1$  e  $R_2$  ditas **compatíveis**
  - $\text{grau}(R_1) = \text{grau}(R_2)$
  - para  $i$  de 1 até  $\text{grau}(R_1)$ :  
 $\text{domínio}(\text{atributo } a_i \text{ de } R_1) = \text{domínio}(\text{atributo } a_i \text{ de } R_2)$
- **Grau do resultado**
  - $\text{grau}(R_1)$  (ou  $\text{grau}(R_2)$ )
- **Nomes dos atributos do resultado**
  - nomes dos atributos da primeira relação ( $R_1$  - relação à esquerda)



# União

- Retorna a união das tuplas de de duas relações  $R_1$  e  $R_2$
- Eliminação automática de duplicatas
- Notação

$\text{relação1} \cup \text{relação2}$

- Exemplo:

$R_1$		
$x$	$y$	$z$
1	1	1
1	2	2
2	2	3
3	1	1

$R_2$		
$x$	$y$	$z$
1	1	1
1	2	1
1	2	3

$R_1 \cup R_2$		
$x$	$y$	$z$
1	1	1
1	2	1
1	2	2
1	2	3
2	2	3
3	1	1

# Diferença

- Retorna as tuplas presentes em  $R_1$  e ausentes em  $R_2$
- Notação

$\text{relaçãol} - \text{relaçãol2}$

- Exemplo:

$R_1$

$x$	$y$	$z$
1	1	1
1	2	2
2	2	3
3	1	1

$R_2$

$x$	$y$	$z$
1	1	1
1	2	1
3	1	1

$R_1 - R_2$

$x$	$y$	$z$
1	2	2
2	2	3

# Intersecção

- Retorna as tuplas comuns a  $R_1$  e  $R_2$
- Notação

$\text{relação1} \cap \text{relação2}$

- Exemplo:

$R_1$

$x$	$y$	$z$
1	1	1
1	2	2
2	2	3
3	1	1

$R_2$

$x$	$y$	$z$
1	1	1
1	2	1
3	1	1

$R_1 \cap R_2$

$x$	$y$	$z$
1	1	1
3	1	1

# Para resolver (usando $\cup$ , $—$ ou $\cap$ )

- 1) buscar o nome e CPF dos médicos e dos pacientes cadastrados no hospital
- 2) buscar o nome, CPF e idade dos médicos, pacientes e funcionários que residem em Florianópolis
- 3) buscar o nome e CPF dos funcionários que recebem salários abaixo de R\$ 500,00 e não estão internados como pacientes
- 4) buscar o número dos ambulatórios onde nenhum médico dá atendimento
- 5) buscar o nome e CPF dos funcionários de Florianópolis que estão internados como pacientes
- 6) buscar o nome e CPF dos médicos pediatras que não atendem nos ambulatórios 101 e 102, e estão internados como pacientes sofrendo de gastrite

# Junção (*Join*)

- Retorna a combinação de tuplas de duas relações  $R_1$  e  $R_2$  que satisfazem um predicado
- Notação

relação1  $\theta$  X relação2

- Exemplo:

$R_1$			$R_2$		$R_1 \theta X R_2 \quad \theta = \sigma_{R1.y > R2.y}$				
$x$	$y$	$z$	$w$	$y$	$x$	$R_1.y$	$z$	$w$	$R_2.y$
1	1	1	1	1	2	2	2	1	1
2	2	2	2	2	3	3	3	1	1
3	3	3			3	3	3	2	2

# Junção Natural (*natural join*)

- Junção na qual  $\theta$  é uma igualdade predefinida entre todos os atributos de mesmo nome presentes em duas relações  $R_1$  e  $R_2$  (*atributos de junção*). Estes atributos só aparecem uma vez no resultado
- Notação

$relaçãol \bowtie relaçãor$

- Derivação

$$R_1 \bowtie R_2 = \pi_{A1, \dots, An, B1, \dots, Bm, \overbrace{C1, \dots, Cx}^{\text{atributos de junção}}} (R_1 \theta X R_2)$$
$$\theta = \sigma_{R1.C1 = R2.C1 \wedge \dots \wedge R1.Cx = R2.Cx}$$

# Junção Natural

- Exemplos

$R_1$

$x$	$y$	$z$
1	1	1
1	1	2
2	2	3

$R_2$

$w$	$y$
1	1
2	2

$R_1 \bowtie R_2$

$x$	$y$	$z$	$w$
1	1	1	1
1	1	2	1
2	2	3	2

$R_1$

$x$	$y$	$z$
1	1	1
1	1	2
2	2	3

$R_2$

$x$	$y$	$w$
1	1	3
2	2	2

$R_1 \bowtie R_2$

$x$	$y$	$z$	$w$
1	1	1	3
1	1	2	3
2	2	3	2

# Junção Natural

- Exemplos

$R_1$

$x$	$y$	$z$
1	1	1
1	1	2

$R_2$

$w$	$t$
1	1
2	2

$R_1 \bowtie R_2$

$x$	$y$	$z$	$w$	$t$
1	1	1	1	1
1	1	1	2	2
1	1	2	1	1
1	1	2	2	2



# Para resolver

- 1) buscar o **número** e a **capacidade** dos ambulatorios do quinto andar e o **nome** dos médicos que atendem neles
- 2) buscar o **nome** e o **salário** dos funcionários de Florianópolis e Palhoça que estão internados como pacientes e têm consulta marcada em 20/10/2006
- 3) buscar o **número** e o **andar** dos ambulatorios onde nenhum médico dá atendimento
- 4) buscar o **número** dos ambulatorios que estão no mesmo andar do ambulatorio 101 e possuem capacidade superior à capacidade dele

# Junções Externas (*outer joins*)

- Junção na qual as tuplas de uma ou ambas as relações que não são combinadas são mesmo assim preservadas no resultado
- Três tipos (exemplos com junção natural)
  - junção externa à esquerda (*left [outer] join*)
    - tuplas da relação à esquerda são preservadas
    - notação:  $\text{rela\c{c}\~ao1} \bowtie \text{rela\c{c}\~ao2}$
  - junção externa à direita (*right [outer] join*)
    - tuplas da relação à direita são preservadas
    - notação:  $\text{rela\c{c}\~ao1} \bowtie \text{rela\c{c}\~ao2}$
  - junção externa completa (*full [outer] join*)
    - tuplas de ambas as relações são preservadas
    - notação:  $\text{rela\c{c}\~ao1} \bowtie \text{rela\c{c}\~ao2}$

# Junções Externas (*outer joins*)

- Exemplos

$R_1$

$x$	$y$	$z$
1	1	1
2	1	2
3	3	3
5	5	5

$R_2$

$x$	$a$	$b$
1	1	1
2	1	2
4	4	4

$R_1 \bowtie R_2$

$x$	$y$	$z$	$a$	$b$
1	1	1	1	1
2	1	2	1	2
3	3	3		
5	5	5		

$R_1 \ltimes R_2$

$x$	$y$	$z$	$a$	$b$
1	1	1	1	1
2	1	2	1	2
4			4	4

$R_1 \ltimes R_2$

$x$	$y$	$z$	$a$	$b$
1	1	1	1	1
2	1	2	1	2
3	3	3		
5	5	5		
4			4	4

# Para resolver

- 1) buscar os dados de todos os médicos e, para aqueles que têm consultas marcadas, mostrar os dados de suas consultas
- 2) buscar os **números** de todos os ambulatorios e, para aqueles ambulatorios nos quais médicos dão atendimento, exibir o **código** e o **nome** dos médicos associados
- 3) mostrar em uma relação o **CPF** e **nome** de todos os pacientes e de todos os médicos, apresentando estes dados de forma relacionada para aqueles que possuem consultas marcadas

# Divisão

- Considera duas relações
  - dividendo (*grau  $m + n$* )
  - divisor (*grau  $n$* )
- *Grau “n”*
  - atributos de mesmo nome em ambas as relações
- Quociente
  - grau “m”
  - atributos da relação dividendo cujos valores associam-se com *todos* os valores da relação divisor
- Notação

*relação1  $\div$  relação2*

# Divisão

- Exemplos

$$R_1$$

$x$	$y$	$z$
1	1	1
1	2	1
2	1	1
2	2	2
3	1	3

$$R_{2a}$$

$z$
1

$$R_{2b}$$

$y$	$z$
1	1

$$R_{2c}$$

$y$
1
2

$$R_1 \overset{\circ}{\underset{\circ}{-}} R_{2a}$$

$x$	$y$
1	1
1	2
2	1

$$R_1 \overset{\circ}{\underset{\circ}{-}} R_{2b}$$

$x$
1
2

$$R_1 \overset{\circ}{\underset{\circ}{-}} R_{2c}$$

$x$	$z$
1	1

# Para resolver

- 1) buscar o **código** dos pacientes que têm consultas marcadas com todos os médicos
- 2) buscar o **nome** e o **CPF** dos médicos que têm consultas marcadas com todos os pacientes
- 3) buscar o **nome** e o **CPF** dos pacientes que têm consultas marcadas com todos os médicos ortopedistas que atendem nos ambulatórios do primeiro andar
- 4) todos os médicos ortopedistas dão atendimento no mesmo ambulatório? Em caso afirmativo, buscar o **número** e o **andar** deste ambulatório

# Atualização de Relações

- Exclusão

- notação

- $\text{relação} \leftarrow \text{relação} - \text{expressãoConsulta}$
    - $\text{relação} \leftarrow \text{expressãoConsulta}$
    - $\text{expressãoConsulta}$  **envolve** relação

- Inclusão

- notação

- $\text{relação} \leftarrow \text{relação} \cup \text{Expr}$
    - $\text{Expr}$  : conjunto de tuplas

- Alteração

- notação

- $\delta_{\{\text{nome\_atributo} \leftarrow \text{Expr}\}}(\text{relação})$
    - $\text{Expr}$  : expressão aritmética ou valor constante



# Atualização de Relações

- Exemplos

$R_1$			$R_2$		
$x$	$y$	$z$	$w$	$t$	$v$
1	1	1	1	3	1
2	1	3	2	2	2
			3	2	3

1) a)  $R_1 \leftarrow R_1 - \sigma_{x=1}(R_1)$

b)  $R_2 \leftarrow \sigma_{t=2}(R_2)$

2)

a)  $R_1 \leftarrow R_1 \cup \{(1,2,2), (1,2,3)\}$

b)  $temp \leftarrow \pi_w(\sigma_{t=2}(R_2))$

$R_1 \leftarrow R_1 \cup (temp \times \{(3,3)\})$

3)

a)  $\delta_{x \leftarrow x+1}(R_1)$

b)  $temp \leftarrow \sigma_{t=2}(R_2)$

$R_2 \leftarrow R_2 - temp$

$\delta_{w \leftarrow w-1}(temp)$

$R_2 \leftarrow R_2 \cup temp$