

Universidade Federal de São Paulo
Bacharelado em Ciência da Computação

Banco de Dados

Álgebra Relacional

(musa@unifesp.br)

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

- Álgebra desenvolvida para descrever operações sobre uma base de dados relacional
- O conjunto de objetos são as Relações
- 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

Álgebra Relacional

- Porque aprender:
 - Compreendendo álgebra relacional é mais fácil aprender SQL
 - Não há SGBD que implementa álgebra diretamente como DML (Data Manipulation Language), mas SQL incorpora cada vez mais conceitos de álgebra
 - Algoritmos de otimização de consulta definidos sobre álgebra (possível uso internamente no SGBD)

Álgebra Relacional

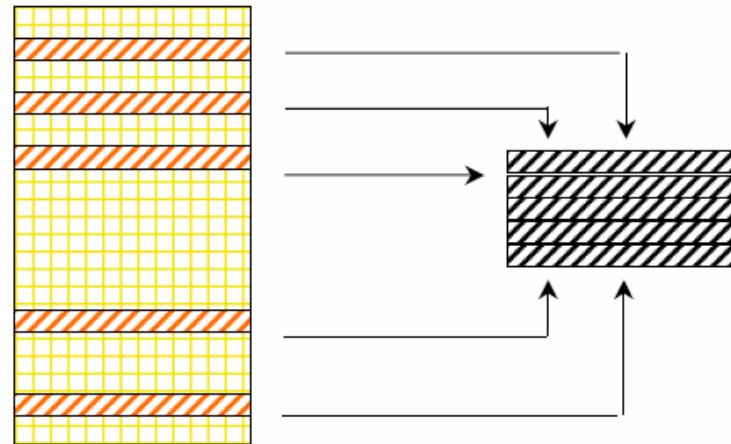
- Operadores sobre conjuntos (uma tabela é um conjunto de linhas):
 - União
 - Interseção
 - Diferença
 - Produto Cartesiano
- Operadores específicos da álgebra relacional:
 - Seleção
 - Projeção
 - Junção
 - Divisão
 - Renomeação

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 (σ)

- Retorna **tuplas** que satisfazem um **predicado**
- A Seleção tem como operando uma relação. O resultado é uma relação que contém as tuplas que obedecem a um determinado critério.



Seleção (σ)

- Sintaxe:

σ <critério de seleção> (<relação>)

onde:

- <relação> é o nome de uma tabela ou uma expressão de álgebra relacional que resulta em uma tabela
- <critério de seleção> é uma expressão booleana que envolve literais e valores de atributos da tabela
- O resultado da seleção tem colunas com os mesmos nomes e domínios da relação de entrada

Seleção (σ) - Exemplo

- Buscar os dados dos pacientes que estão com sarampo

σ doença="sarampo"(Pacientes)

Seleção (σ) - Exemplo

- Buscar os dados dos médicos ortopedistas com mais de 55 anos
 σ especialidade="ortopedista" and idade ≥ 55 (Medicos)

Projeção (π)

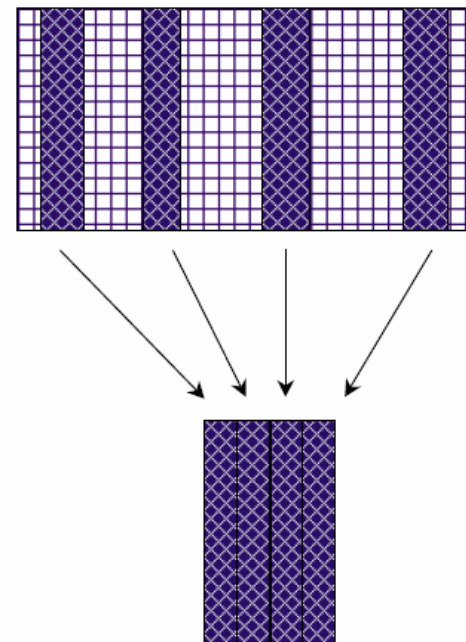
- Retorna um ou mais **atributos** de interesse
- A Projeção tem como operando uma relação. O resultado é uma relação que contém apenas as colunas selecionadas.

R

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

$\pi_{x,y} R$

x	y
1	1
2	2



* Elimina duplicatas

Projeção (π)

- Sintaxe:

π <lista de atributos> (<relação>)

onde:

- <relação> é o nome de uma tabela ou uma expressão de álgebra relacional que resulta em uma tabela.
- <lista de atributos> é uma lista que contém nomes de colunas da tabela operando.

Projeção (π) – Exemplo

- Buscar o nome e a especialidade de todos os médicos

π nome, especialidade (Medico)

Projeção e Seleção

- Operadores diferentes podem ser aninhados
- Exemplo:
 - buscar o número dos ambulatorios do terceiro andar

$\pi_{\text{nroa}} \sigma (\text{andar}=3) (\text{Ambulatorios})$

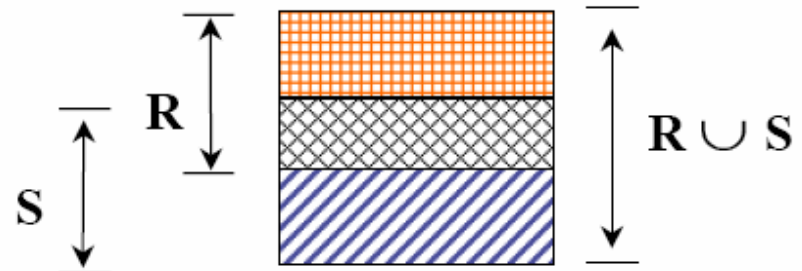
Operações - Teoria dos Conjuntos

- A álgebra relacional utiliza da teoria de conjuntos quatro operadores: **União, Intersecção, Diferença e Produto Cartesiano**
- Nos três casos, a operação possui duas tabelas como operando.
- As tabelas devem ser compatíveis:
 - possuir o mesmo número de colunas
 - o domínio da i -ésima coluna de uma tabela deve ser idêntico ao domínio da i -ésima coluna da outra.
- Quando os nomes das colunas forem diferentes, adota-se a convenção de usar os nomes das colunas da primeira tabela.

União (\cup)

- Retorna a união das tuplas de duas relações R e S
- Eliminação automática de duplicatas
- Notação:
 - $R \cup S$

R			S			$R \cup S$		
x	y	z	x	y	z	x	y	z
1	1	1	1	1	1	1	1	1
1	2	2	1	2	1	1	2	1
2	2	3	1	2	3	1	2	2
3	1	1				1	2	3
						2	2	3
						3	1	1



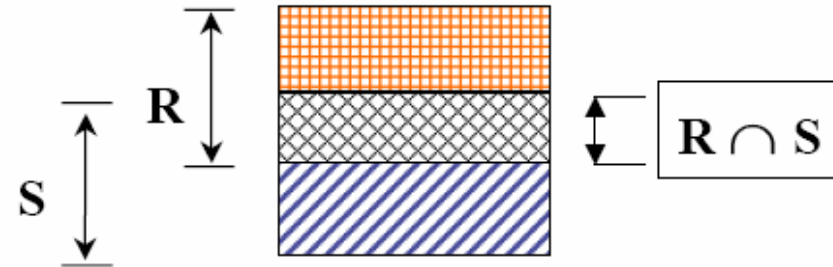
União (\cup) - Exemplo

- buscar o nome e CPF dos médicos e dos pacientes cadastrados no hospital

$$\pi_{(\text{nome}, \text{CPF})}(\text{Medicos}) \cup \pi_{(\text{nome}, \text{CPF})}(\text{Pacientes})$$

Intersecção (\cap)

- Retorna as tuplas comuns a R e S
- Notação:
 - $R \cap S$



R			S			$R \cap S$		
x	y	z	x	y	z	x	y	z
1	1	1	1	1	1	1	1	1
1	2	2	1	2	1			
2	2	3	3	1	1	3	1	1
3	1	1						

Intersecção (\cap) - Exemplo

- buscar o nome e CPF dos funcionários de Porto Alegre que estão internados como pacientes

$\pi_{(\text{nome}, \text{CPF})} \sigma_{\text{cidade}=\text{"Porto Alegre"}}(\text{Funcionarios})$

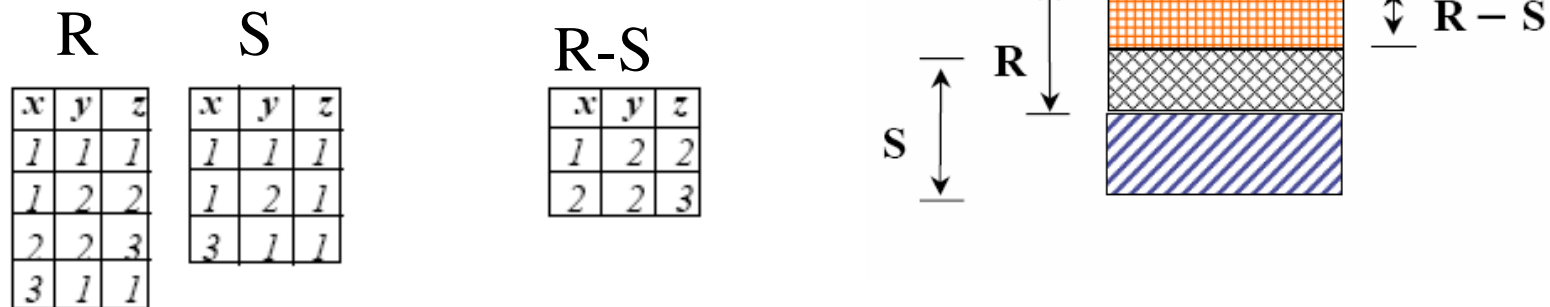
\cap

$\pi_{(\text{nome}, \text{CPF})} (\text{Pacientes})$

Diferença (-)

- Retorna as tuplas presentes em R e ausentes em S
- Notação:

$R - S$



Diferença (-) - Exemplo

- buscar o número dos ambulatórios onde nenhum médico dá atendimento

$$\pi_{(\text{nroa})}(\text{Ambulatorio}) - \pi_{(\text{nroa})}(\text{Medicos})$$

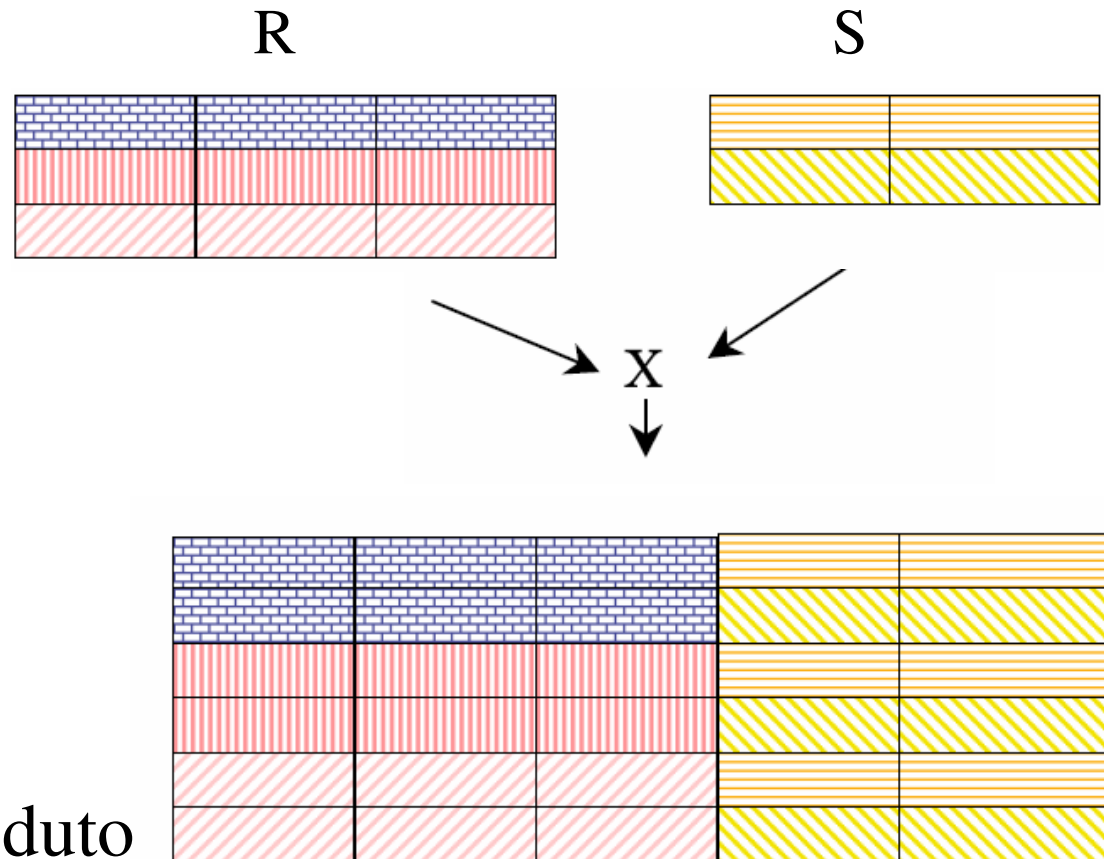
Produto Cartesiano (x)

- Retorna todas as combinações de tuplas de duas relações R e S
- O resultado é uma relação cujas tuplas são a combinação das tuplas das relações R e S, tomando-se uma tupla de R e concatenando-a com uma tupla de S
- Notação:
 - $R \times S$

Produto Cartesiano (x)

Total de atributos do produto cartesiano =
num. atributos de R +
num. atributos de S

Número de tuplas do produto cartesiano = num. tuplas de R x
num tuplas de S



Produto Cartesiano (x)

- Exemplo:

R			S	
x	y	z	w	y
1	1	1	1	1
2	2	2	2	2
3	3	3		

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

Produto Cartesiano - Exemplo

- buscar o nome dos médicos que têm consulta marcada e as datas das suas consultas

π medicos.nome, consultas.data σ medicos.codm=consultas.codm (Medicos)
x (Consultas)

Produto Cartesiano - Exemplo

- buscar, para as consultas marcadas para o período da manhã (7hs-12hs), o nome do médico, o nome do paciente e a data da consulta

Produto Cartesiano - Exemplo

- buscar, para as consultas marcadas para o período da manhã (7hs-12hs), o nome do médico, o nome do paciente e a data da consulta

π medicos.nome, pacientes.nome, consultas.data

σ (consultas.hora \geq 7 and consultas.hora \leq 12) and

medicos.codm=consultas.codm and consultas.codp=pacientes.codp

(Medico) x (Consultas) x (Pacientes)

Exercícios

- buscar os dados dos ambulatórios do quarto andar. Estes ambulatórios devem ter capacidade igual a 50 ou o número do ambulatório deve ser superior a 10
- buscar os números dos ambulatórios que os médicos psiquiatras atendem
- buscar o nome e o salário dos funcionários de Jacareí e Taubaté que estão internados como pacientes e têm consulta marcada com ortopedistas
- buscar os CPF dos funcionários que não são pacientes
- Buscar os CPF dos funcionários que são pacientes

Junção

- Retorna a combinação de tuplas de duas relações R e S que satisfazem um predicado
- **Seleção** combinada com **Produto Cartesiano**
- Sintaxe:
 - $\langle \text{Relação S} \rangle \bowtie \langle \text{critério} \rangle \langle \text{Relação R} \rangle$
 - onde:
 - $\langle \text{relação} \rangle$ é o nome de uma tabela ou uma expressão de álgebra relacional que resulta em uma tabela
 - $\langle \text{critério} \rangle$ é uma expressão booleana envolvendo literais e valores de atributos das duas tabelas.

Junção - Exemplo

- buscar o número dos ambulatorios e o nome dos médicos que atendem neles

π Ambulatorio \bowtie (ambulatorio.nroa=medicos.nroa) Medicos
ambulatorio.nroa, medicos.nome

Junção Natural

- Junção na qual \bowtie é uma igualdade predefinida entre todos os atributos de mesmo nome presentes em duas relações R e S (atributos de junção). Estes atributos só aparecem uma vez no resultado
- Notação:
 - $\langle \text{relação} \rangle R \theta \langle \text{relação} \rangle S$

Onde:

$\langle \text{relação} \rangle$ é o nome de uma tabela ou uma expressão de álgebra relacional que resulta em uma tabela

Junção Natural - Exemplo

Exemplo: buscar o número dos ambulatórios e o nome dos médicos que atendem neles

π Ambulatorio θ Medicos
ambulatorio.nroa, medicos.nome

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
- Tipos:
 - junção externa à esquerda (left [outer] join)
 - junção externa à direita (right [outer] join)
 - junção externa completa (full [outer] join)

Junção Externa à esquerda

- tuplas da relação à esquerda são preservadas
- Notação:
 - $R \bowtie S$

R

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

S

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

$R \bowtie S$

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

Junção Externa à esquerda

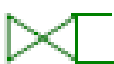
- Exemplo: buscar os dados de todos os médicos e para aqueles que têm consultas marcadas, mostrar os dados de suas consultas

Médicos  Consulta
(medicos.codm=consulta.codm)

Junção externa à direita

- tuplas da relação à direita são preservadas

- Notação

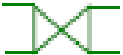
R  S

R			S		
<i>x</i>	<i>y</i>	<i>z</i>	<i>x</i>	<i>a</i>	<i>b</i>
1	1	1	1	1	1
2	1	2	2	1	2
3	3	3	4	4	4
5	5	5			

R			S		
<i>x</i>	<i>y</i>	<i>z</i>	<i>a</i>	<i>b</i>	
1	1	1	1	1	
2	1	2	1	2	
4			4	4	

Junção externa completa

- tuplas de ambas as relações são preservadas


- Notação:
R  S

R

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

S

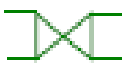
x	a	b
1	1	1
2	1	2
4	4	4

R  S

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

Junção externa completa

- Exemplo: mostrar em uma relação o CPF e nome de todos os pacientes e todos os funcionários apresentando estes dados de forma relacionada quando existente

Pacientes  Funcionários (paciente.cpf=funcionarios.cpf)

π paciente.cpf, paciente.nome

Junção externa completa

- Exemplo: 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

π paciente.CPF, pacientes.nome

medicos.codm=consultas.codm and
consultas.codp=pacientes.codp

(Paciente) \bowtie (Consultas) \bowtie (Medicos)

Divisão

- Útil para responder questões como: “encontre os pacientes que consultaram com todos os médicos”
- Sintaxe: $R : S$
- Os nomes das colunas de S devem estar contidos em R
- A relação resultante tem como nomes das colunas aqueles que aparecem em R mas não aparecem em S ($R-S$)
- Para que uma linha apareça no resultado, é necessário que a sua concatenação com cada linha de R apareça também em S .

Divisão

R

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

S

A
a1
a2
a3

T

B
B1
B4

Divisão

R

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

S

A
a1
a2
a3

T

Resultado

B
b1
b4

Divisão

- buscar o código dos pacientes que têm consultas marcadas com todos os médicos

π Consultas : Médicos

codm,codp

codm

Resposta: CODP

CONSULTA

Codm	Codp
1	1
2	1
1	2
1	3
2	3

MEDICO

Codm
1
2

Codp
1
3