

Base de Dados

04 - Álgebra Relacional

Vitor Vaz da Silva

Base de Dados - Índice

- Álgebra Relacional

Álgebra

É um **sistema matemático** constituído por:

Operandos – variáveis, valores ou expressões

Operadores – símbolos que
representam procedimentos que
permitem obter valores
a partir de outros valores

Álgebra Relacional

Operandos

Relações

Valores que representam Relações

Operadores

Traduzem as relações mais simples e comuns a realizar sobre Relações

Resulta numa álgebra a utilizar para efectuar pesquisas sobre relações

Álgebra Relacional

Query Language: (Linguagem de Pesquisa)

Permite manipular e retirar dados
de uma Base de Dados

Não é uma Linguagem de Programação

Tem uma forte componente formal baseada em
lógica relacional

Álgebra Relacional

Álgebra Relacional:

Operacional (procedimental)

Representa planos de execução

Cálculo Relacional: (não operacional)

Descreve

o que se quer e

não como se quer feito : Declarativa

Base de Dados

Álgebra Relacional

Subconjunto

União

Intersecção

Cardinalidade

Número de elementos

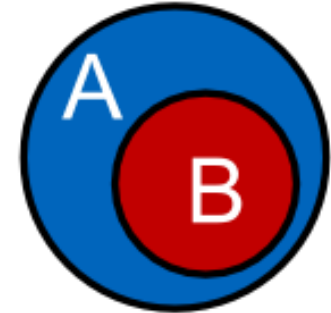
Produto Cartesiano

$A \times B \times C \times D \{(a_1, b_1, c_1, d_1), (a_1, b_1, c_1, d_2), \dots\}$

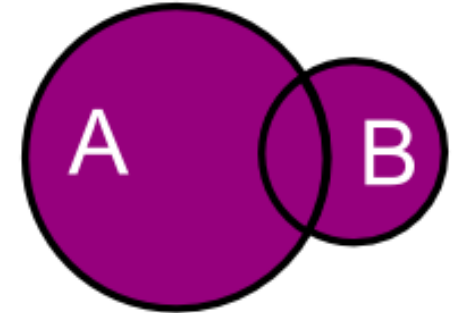
Grau

Dimensão do par ordenado (a, b, c) é 3

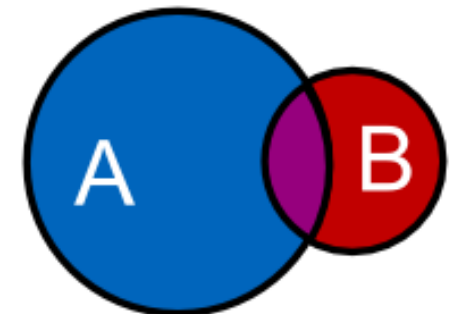
$$B \subseteq A$$



$$A \cup B$$

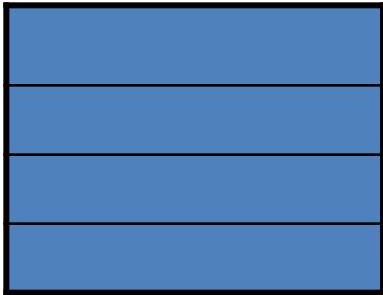


$$A \cap B$$

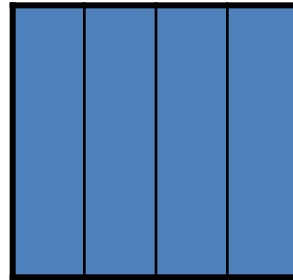


Representação Gráfica

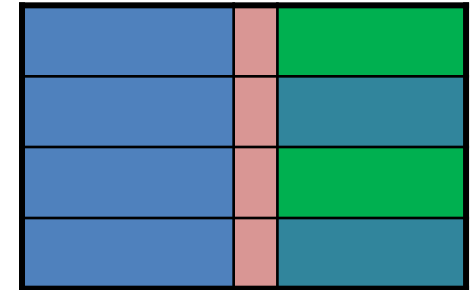
Seleção



Projeção



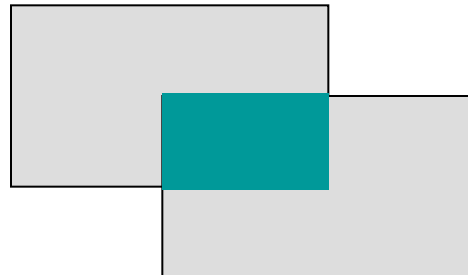
Junção



União



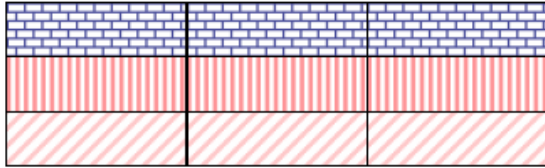
Intersecção



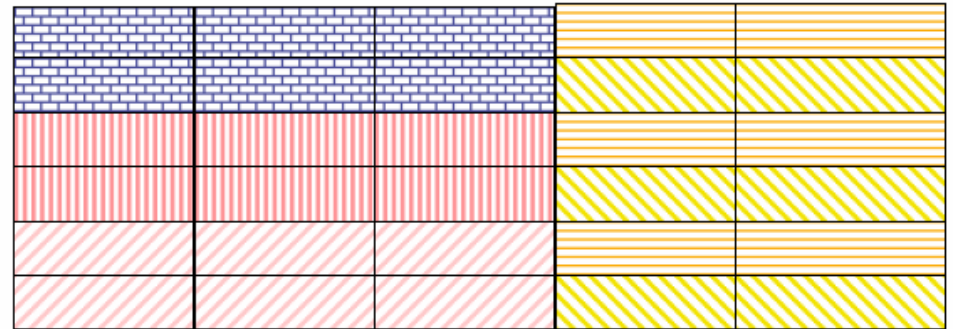
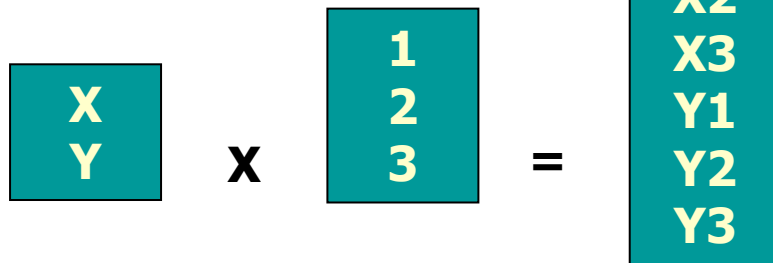
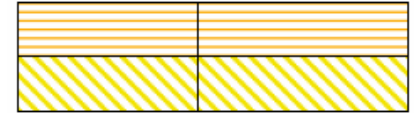
Diferença



Representação Gráfica



Produto Cartesiano



Representação Gráfica

Divisão

X1
X2
X3
Y1
Y2
Y3

÷

X
Y

=

1
2
3

Representação Gráfica

Blue brick	Blue brick	Blue brick	Orange horizontal	Orange horizontal
Blue brick	Blue brick	Blue brick	Yellow diagonal	Yellow diagonal
Red vertical	Red vertical	Red vertical	Orange horizontal	Orange horizontal
Red vertical	Red vertical	Red vertical	Yellow diagonal	Yellow diagonal
Pink diagonal	Pink diagonal	Pink diagonal	Orange horizontal	Orange horizontal
Pink diagonal	Pink diagonal	Pink diagonal	Yellow diagonal	Yellow diagonal

Divisão

Blue brick	Blue brick	Blue brick
Red vertical	Red vertical	Red vertical
Pink diagonal	Pink diagonal	Pink diagonal

X1
X3
Y1
Y2
Y4
X2
X4

÷

X
Y

=

1
2

Orange horizontal	Orange horizontal
Yellow diagonal	Yellow diagonal

Representação Gráfica

Join

idx	par
a	x
b	y
c	y

par	res
x	asd
y	qwe
z	zxc

idx	par	res
a	asd	x
b	y	qwe
c	y	qwe

Representação Gráfica

Renomear

A	B	C

A	D	C

Representação Gráfica

Agregar

A	B	C

Conta
Máximo
Mínimo
Soma
Média



*Produz **um** único valor*

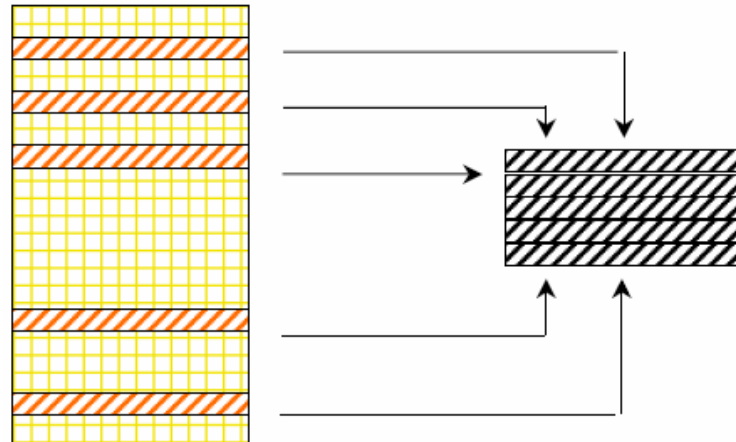
Simbologia

Seleção	$\sigma_F(R)$	Selection
Projeção	$\Pi_{i_1, i_2, \dots, i_n}(R)$	Projection
Junção	$R \bowtie S$	Join
União	$R \cup S$	Union
Intersecção	$R \cap S$	Intersect
Diferença	$R - S$	Difference
Produto Cartesiano	$R \times S$	Product
Divisão	$R \div S$	Division
Renomear	$\rho_T(R)$	Renaming
Agregar	$\mathcal{F}(R)$	Agregation

Seleção (σ)

Retorna tuplos que satisfazem uma condição
Age como um filtro.

Possui os mesmos atributos da relação de entrada



Seleção (σ)

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

Sigma σ é o símbolo que representa a seleção

$\langle \text{condição de seleção} \rangle$ é uma expressão booleana que envolve literais e valores de atributos da relação

CLAUSULAS: $\langle \text{nome do atributo} \rangle \langle \text{operador relacional} \rangle$
 $\{ \langle \text{valor constante} \rangle \mid \langle \text{nome do atributo} \rangle \}$

Nome do atributo: é um atributo de $\langle R \rangle$

Operador relacional: =, <, <=, >, >=, <>

Valor constante: é um valor do domínio do atributo

Podem ser ligadas pelos operadores AND, OR e NOT

$\langle R \rangle$ é o nome de uma relação ou uma *expressão da álgebra relacional* de onde se encontram os tuplos

Seleção (σ)

$$R1 := \sigma_C(R2)$$

C é a condição (como num "if") referente aos atributos de R2

R1 são todos os tópicos de R2 que satisfazem a condição C

Relação

Relação

Atributos

Venda

bar	bebida	preço
O’Gillans	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Tuplo *ou*
Par Ordenado

Seleção (σ)

Venda

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Como seleccionar as linhas em que o bar é a “Adega do Zé” ?

Seleção (σ)

Venda

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

$\text{DoZe} := \sigma_{\text{bar}=\text{"Adega do Zé"}}(\text{Venda})$

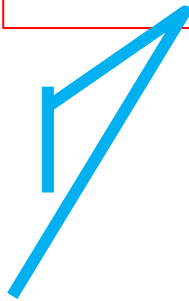


Seleção (σ)

$\text{DoZe} := \sigma_{\text{bar}=\text{"Adega do Zé"}}(\text{Venda})$

Relação

DoZe



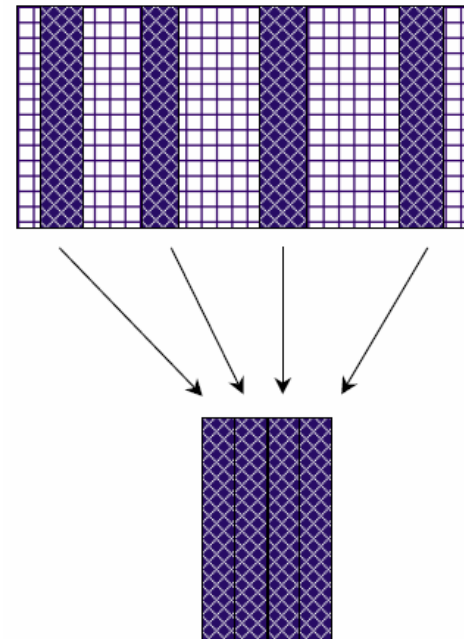
bar	bebida	preço
Adega do Zé	Guinness	2.50
Adega do Zé	Sagres	1.75

Projeção (π)

Retorna um ou mais atributos de interesse

O resultado é uma relação que contém apenas as colunas seleccionadas.

Elimina duplicados



Projeção (π)

Sintaxe:

π <lista de atributos> (<R>)

<lista de atributos> é uma lista que contém nomes de colunas de uma ou mais relações
(a ordem das colunas é a definida na lista)

<R> é o nome da relação
ou uma expressão da álgebra relacional de onde a lista de atributos será retirada

Projeção (π)

Venda

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Quais as bebidas e preços praticados?

Projeção (π)

Venda

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Preços := $\pi_{\text{bebida, preço}}(\text{Venda})$



Projeção (π)

$\text{Preços} := \pi_{\text{bebida, preço}}(\text{Venda})$

Venda

bebida	preço
Guinness	2.50
Sagres	1.75
Tuborg	2.20
Sagres	1.50

Projeção Extendida (π)

Sintaxe:

π <lista de atributos> (<R>)

<lista de atributos> pode conter:
expressões aritméticas: $a+b \rightarrow c$
repetir atributos: a, a

<R> Uma Relação, ex: $R(a, b)$

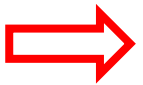
$\pi_{a+b \rightarrow c, a, a} R$

Projeção Extendida (π)

Venda

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Promo := $\pi_{\text{bebida, preço*0.3} \rightarrow \text{prom, bebida}}(\text{Venda})$



Projeção Extendida (π)

Promo := $\pi_{\text{bebida, preço} \cdot 0.3 \rightarrow \text{prom, bebida}}$ (Venda)

Promo

As colunas podem ser repetidas mas os títulos têm de ser distintos (numeração automática)

bebida1	prom	bebida2
Guinness	0.75	Guinness
Sagres	0.53	Sagres
Tuborg	0.66	Tuborg
Sagres	0.45	Sagres
Guinness	0.75	Guinness

Esquema

- O esquema de uma relação consiste na sua estrutura:

o nome e o tipo de dados dos atributos

- Duas relações têm o mesmo esquema quando a estrutura é a mesma
(a ordem dos atributos não interfere)

Intersecção ($\cap$)

Requer que as duas relações fornecidas como argumento tenham o mesmo esquema.

(ex: $R \cap S$)

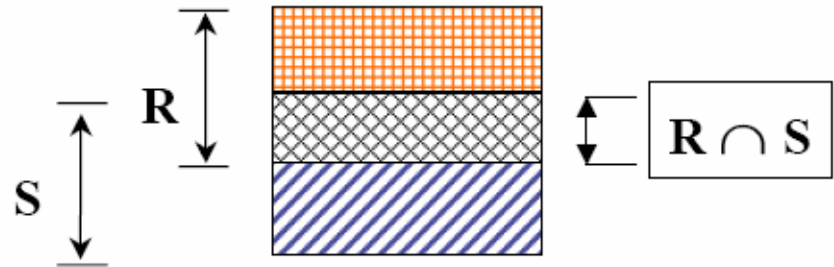
Resulta numa nova relação
com o mesmo esquema
com os tuplos comuns a R e S

(pode haver tuplos repetidos!)

Intersecção ($\cap$)

Resulta numa nova relação com o mesmo esquema e com os tuplos comuns aos termos da operação

Notação: $R \cap S$



Intersecção ($\cap$)

R

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

$\cap$

S

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

$R \cap S =$

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

União ($\cup$)

Requer que as duas relações fornecidas como argumento tenham o mesmo esquema.

Resulta em uma nova relação,
com o mesmo esquema,
cujo conjunto de linhas é a união dos conjuntos de
linhas das relações dadas como argumento.

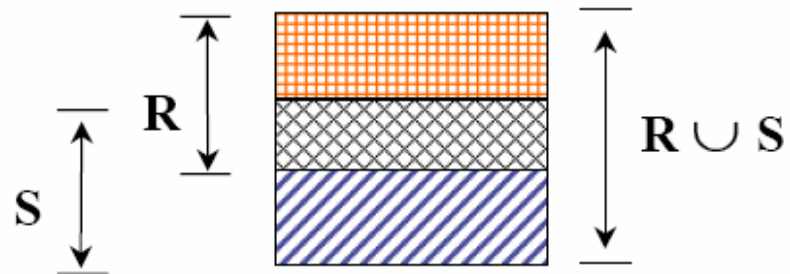
Eliminação automática de duplicados

União ($\cup$)

Notação:

$$R \cup S$$

Retorna a união dos tuplos de duas relações R e S



União ($\cup$)

R

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

S

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

$\cup$

$R \cup S =$

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

Diferença (-)

Requer que as duas relações fornecidas como argumento tenham o mesmo esquema.

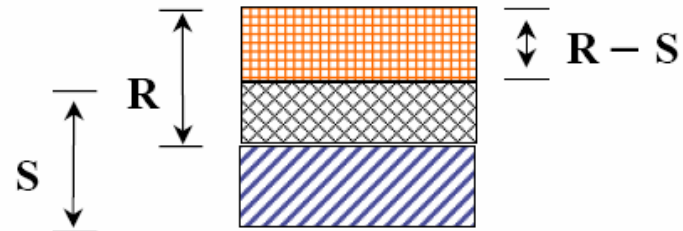
Resulta numa nova relação,
com o mesmo esquema, cujo
conjunto de linhas é o
conjunto de linhas da primeira relação
menos o
conjunto de linhas existentes na segunda.

Diferença (-)

Notação:

$R - S$

Retorna os tuplos
presentes em R e
ausentes em S



Diferença (-)

R

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

S

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

-

R - S =

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

Renomear(ρ)

- Sintaxe:

ρ **<N>**(**<lista de atributos>**) (**<R>**)

- **<lista de atributos>** é uma lista que contém os novos nomes das colunas
- **<N>** é o nome da nova relação
- **<R>** é o nome da relação de onde todos os tuplos vão ser copiados

Renomear (ρ)

O operador ρ dá um novo esquema a uma relação

$$\rho_{R1(A_1, \dots, A_n)}(R2)$$

cria a relação R1 com os atributos $A_1, \dots, A_n$
e com os mesmos tuplos que R2.

Notação simplificada: $R1(A_1, \dots, A_n) := R2$

Renomear(ρ)

Venda

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Como trocar o título das colunas de *bar*, *bebida*, *preço* para: *local*, *tipo*, *valor* ?

Renomear(ρ)

Venda

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

ρ Bares (local, tipo, valor) (Venda)



Renomear(ρ)

ρ Bares (local, tipo, valor) (Venda)

Bares

local	tipo	valor
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Produto Cartesiano (x)

Retorna todas as combinações de tuplos de duas relações

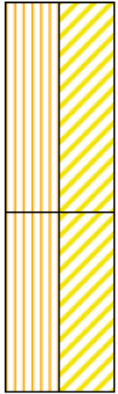
Notação:

$$R \times S$$

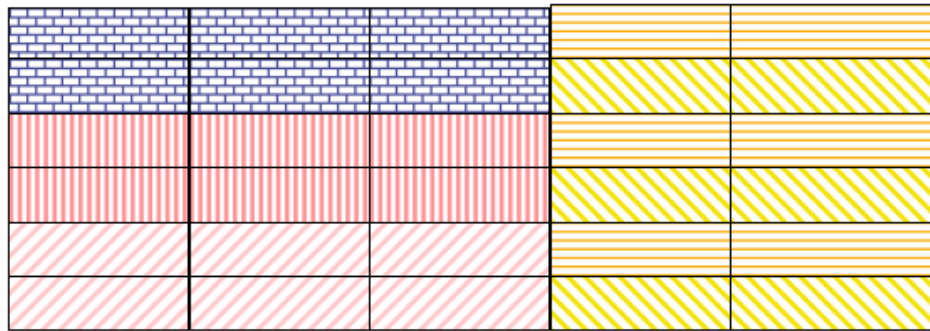
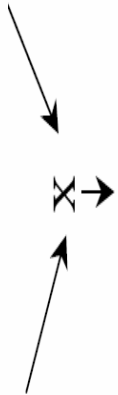
O resultado é uma relação cujos tuplos são a combinação dos tuplos das relações R e S, tomando-se um tuplo de R e concatenando-o com cada um dos tuplos de S

Produto Cartesiano (x)

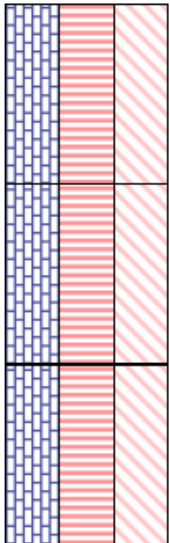
S



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



R



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

Produto Cartesiano (x)

$$R3 := R1 \times R2$$

Emparelha cada tuplo t_1 de $R1$ com cada tuplo t_2 de $R2$

A concatenação $t_1 t_2$ é o novo tuplo de $R3$

O novo esquema tem primeiro os atributos de $R1$ seguidos dos de $R2$

No caso de duplicação de atributos são usados os nomes da relação. ex: $R1.A$ $R2.A$

Produto Cartesiano (x)

R

x	y	t
1	1	1
1	2	2
2	2	3
3	1	1

S

z	t
1	1
1	2

x

R x S =

x	y	R.t	z	S.t
1	1	1	1	1
1	1	1	1	2
1	2	2	1	1
1	2	2	1	2
2	2	3	1	1
2	2	3	1	2
3	1	1	1	1
3	1	1	1	2

Junção($\bowtie$) – [Natural] Join

Junta duas relações através de um atributo com o mesmo nome e acrescentando os demais

Notação

$R3 := R1 \bowtie R2$

Venda **Junção(⋈) – [Natural] Join**

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Local	bar	local
	O'Gilins	Cais do Sodré
	Adega do Zé	Chelas

Junção($\bowtie$) – [Natural] Join

Bares := Venda $\bowtie$ Local

Bares

bar	bebida	preço	local
O'Gilins	Guinness	2.50	Cais do Sodré
Adega do Zé	Sagres	1.75	Chelas
Adega do Zé	Guinness	2.50	Chelas

Em que locais há vendas ? *ou*
As vendas são feitas em que locais ?

Junção Condicional ($\bowtie_c$) – Theta Join

Equivalente à junção com uma condição

Notação:

$$R3 := R1 \bowtie_c R2$$

equivalente a: $R3 := \sigma_c(R1 \times R2)$

Venda $\Join_c$ – Theta Join

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Local

loja	local
O'Gilins	Cais do Sodré
Adega do Zé	Chelas

Junção($\bowtie$) – [Natural] Join

Bares := Venda $\bowtie$ _{Venda.bar=Local.loja} Local

Como não há um atributo comum, é necessária a condição

Bares

bar	bebida	preço	loja	local
O'Gilins	Guinness	2.50	O'Gilins	Cais do Sodré
Adega do Zé	Sagres	1.75	Adega do Zé	Chelas
Adega do Zé	Guinness	2.50	Adega do Zé	Chelas

Funções de Agregação ($\mathcal{F}$)

Forma geral: $G_1, G_2, \dots, G_i \mathcal{F}_{F_1(A_1), F_2(A_2), \dots, F_n(A_n)}(R)$

R é uma relação

G_i são atributos de R sobre os quais recai a agregação

Cada F_n é a função de agregação aplicada ao atributo A_n

Passos:

A relação é dividida em grupos G_i

As funções de agregação são aplicadas a cada grupo

Funções de Agregação ($\mathcal{F}$)

Quantas pessoas terminaram o puzzle?

Terminou(nome, puzzle)

nome $\mathcal{F}$ **count-distinct**(*puzzle*)(*Terminou*)

Select count(distinct puzzle)

From Terminou

Group by nome

Funções de Agregação ($\mathcal{F}$)

Quantas pessoas terminaram o puzzle mais do que cinco vezes?

Terminou(nome, puzzle)

$\sigma_{\text{count-distinct}(\text{puzzle}) > 5} (\text{nome } \mathcal{F}_{\text{count-distinct}(\text{puzzle})}(\text{Terminou}))$

Select nome, count(distinct puzzle)

From Terminou

Group by nome

Having count(distinct puzzle) > 5

Funções de Agregação ($\mathcal{F}$)

Qual o preço mais alto e a *média* dos preços *por tipo de vinho*?

ρ (tipo_vinho, “maior”, “médio”) ($\mathcal{F}$ tipo_vinho MAX (preço), AVG (preço) (vinho))

```
SELECT tipo_vinho,  
MAX (preço) AS “maior”,  
AVG (preço) AS “médio”  
FROM vinho  
GROUP BY tipo_vinho
```

Como não há clausula Having não se coloca o σ mas como há nomes para as colunas utiliza-se o ρ

NULL

O *null* afecta tudo!

null significa unknown “desconhecido” or “não existe”

É necessário especificar o que acontece quando o null está presente.
Pode ser arbitrário, é preciso ler o manual das especificações

Operações aritméticas (+, −, *, /) que envolvem null têm null como resultado

Ex: $5 + \text{null}$ dá “desconhecido” : *null*

Comparar com null tem como resultado null

Ex: $\text{null} = \text{null}$ dá “desconhecido” : *null*

NULL $\Leftrightarrow$ UNKNOWN

AND

True $\wedge$ unknown = unknown

False $\wedge$ unknown = False

unknown $\wedge$ unknown = unknown

OR

true $\vee$ unknown = true

false $\vee$ unknown = unknown

unknown $\vee$ unknown = unknown

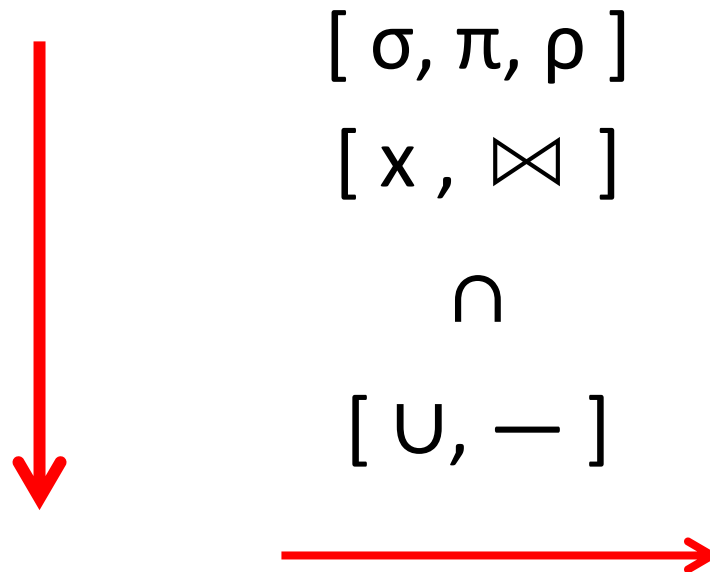
NOT

$\neg$ unknown = unknown

Expressões

Regras de precedência:

Prioridade cima-> baixo e esquerda->direita



Árvores

Uma expressão pode ser escrita em forma de árvore.

*(A árvore não tem ciclos fechados!
Só há um caminho entre dois pontos!)*

Folhas – são operandos (variáveis ou relações)

Nós – são operadores

Árvores

Exemplo:

Utilizando as relações

Local (loja, local)

Venda(bar, bebida, preço)

Procure o nome dos bares
que estão no Cais do Sodré,
ou vendem Sagres a um preço inferior a 1.70

Venda

bar	bebida	preço
O'Gilins	Guinness	2.50
Adega do Zé	Sagres	1.75
Trovador	Tuborg	2.20
Trovador	Sagres	1.50
Adega do Zé	Guinness	2.50

Local

loja	local
O'Gilins	Cais do Sodré
Adega do Zé	Chelas

Árvores

Exemplo: Utilizando as relações

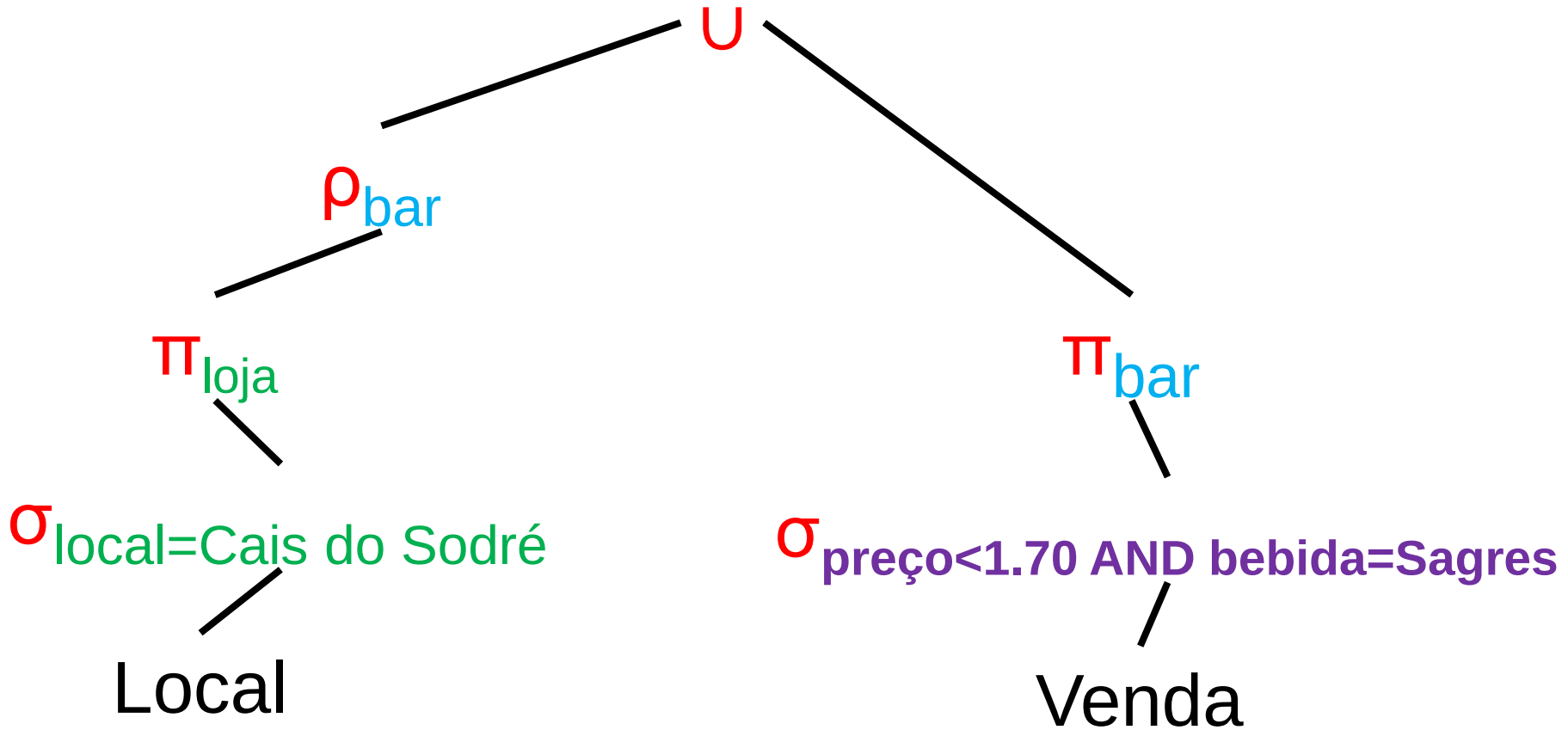
Local (loja, local) e Venda(bar, bebida, preço)
procure o

nome de todos os bares

que estão no Cais do Sodré

ou vendem Sagres a um preço inferior a 1.70

Árvore



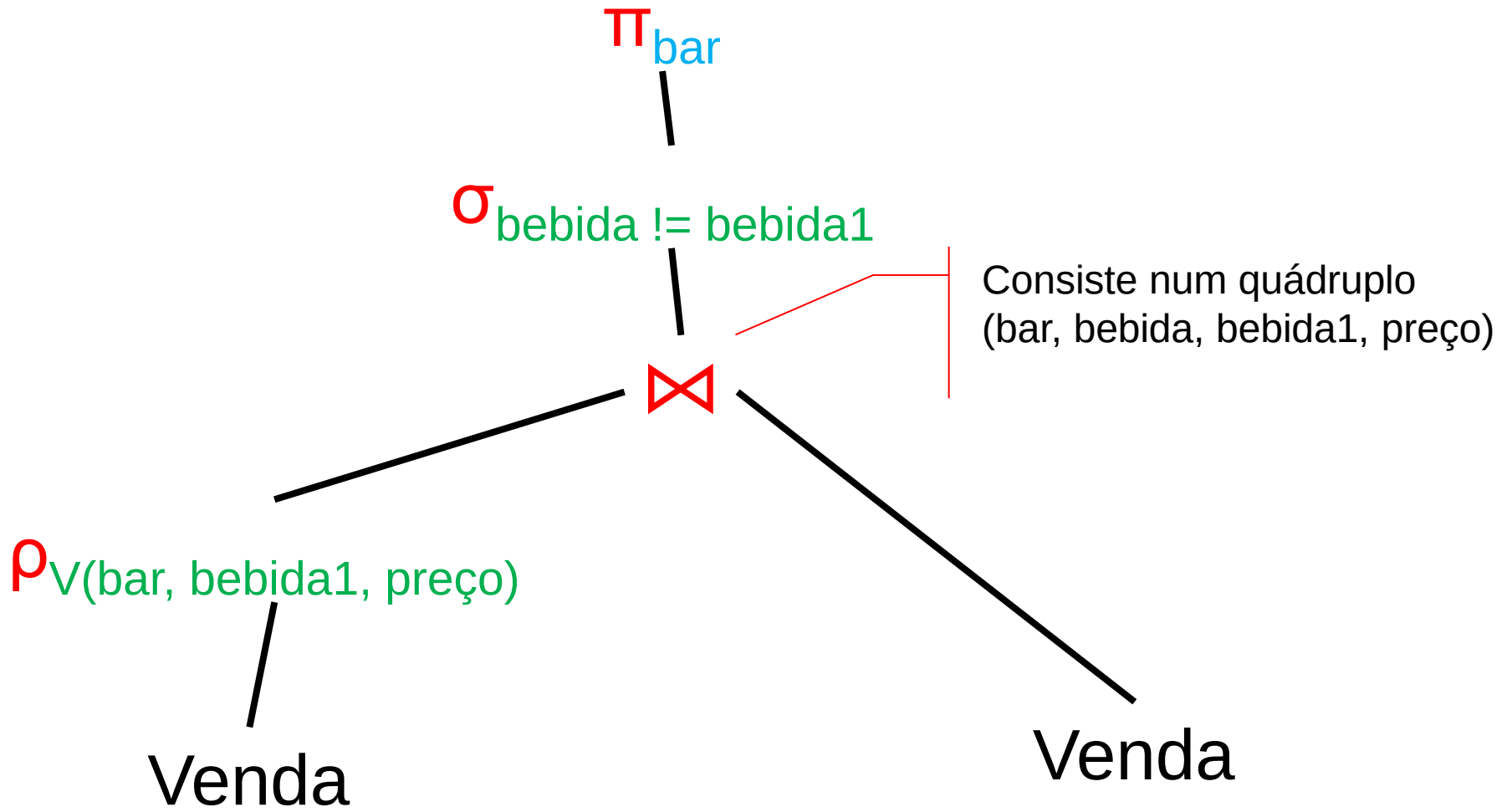
Árvores

Exemplo: Utilizando a relação

Venda(bar, bebida, preço)
procure o

nome de todos os **bares** que
vendem bebidas diferentes ao mesmo preço

Árvore



Resumo - Esquemas

União, Intersecção e Diferença:

Os esquemas dos dois operandos tem de ser igual e o resultado também fica com o mesmo esquema

Seleccção:

O esquema resultante é idêntico ao do operando

Projecção:

O esquema é dado pela lista dos atributos

Resumo - Esquemas

Produto, Theta-join:

O esquema resultante contém os atributos de ambas relações (iguais $\rightarrow$ A1, A2,...)

Natural join:

O esquema resultante é a união dos atributos das duas relações (iguais ficam iguais; o mesmo atributo)

Renomear:

O esquema é indicado pelo operador

Saco – Multi Conjunto

Um Bag (saco) é idêntico a um conjunto mas pode ter elementos repetidos.

Seleccção, Projecção, Produtos e Junções

operam em pares de tóiplos por isso os repetidos são utilizados

Saco – Multi Conjunto

As regras são diferentes.

Com sacos:

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

$$S \cup S \neq S$$

$$\text{ex: } \{1\} \cup \{1\} = \{1, 1\}$$

O conjunto é idempotente $S \cup S = S$; o saco não é.

Calculadora para Algebra Relacional

- <http://dbis-uibk.github.io/relax/help.htm>
- <http://dbis-uibk.github.io/relax/calc.htm#>
- Converte Algebra Relacional de e para SQL

Base de Dados

Referências

- RAMAKRISHNAN, R. - Database Management Systems - McGraw-Hill
- http://www.inf.ufsc.br/~vania/teaching/ine5613/Aula_AlgebraRelacional.ppt
- http://www.inf.furb.br/~arv/bd/Algebra_Relacional.ppt
- <http://infolab.stanford.edu/~ullman/fcdb/aut07/slides/ra.pdf>
- <http://users.cms.caltech.edu/~donnie/cs121/CS121Lec03.pdf>
- <http://dbis-uibk.github.io/relax/>