Bases de Dados



Faculdade de Engenharia

FEUP

 $\begin{pmatrix} 1 \end{pmatrix}$

- INTRODUÇÃO
- MODELOS CONCEPTUAIS
 - Diagrama de Classes UML
 - Modelo Entidade-Associação (E-A)
- MODELO RELACIONAL
- LINGUAGEM DE DEFINIÇÃO DE DADOS
- INTERROGAÇÃO DE DADOS

.ÁLGEBRA RELACIONAL

Linguagem de Manipulação de Dados (LMD)

Observação: baseado em slides desenvolvidos pelo Prof. Gabriel David da FEUP

João Mendes Moreira FEUP

Índice

2

- Linguagens relacionais
- Operações da ágebra relacional
- Leis algébricas
- Linguagem de interrogação
- Extensões à álgebra relacional

João Mendes Moreira FEUP

Linguagens relacionais



- Notações para expressar perguntas:
 - o Algébrica aplicação de operadores a relações
 - o Lógica fórmula que os tuplos da resposta devem satisfazer
- Limitações da álgebra relacional:
 - o independência física dos dados
 - ▼ linguagem só com construções relativas ao modelo de dados (operações sobre relações) que não dependam da implementação
 - o optimização das perguntas
 - ▼ restrição no poder expressivo: sem recursividade (incapaz de computar o fecho transitivo)
 - o relações finitas
 - × Noção de complementar é proibida



Reunião

- 1 Reunião R∪S
 - o é o conjunto dos tuplos que estão em R, em S, ou em ambas
 - ▼ R e S da mesma aridade (mesma aridade = mesmo nº de atributos)
 - ➤ Domínios dos atributos de R e S devem ser compatíveis:
 - o Domínios do 1º atributo de R e S devem ser compatíveis
 - o Domínios do 2º atributo de R e S devem ser compatíveis
 - o etc.

R			S		
Α	В	С	D	E	F
а	b	С	b	g	а
d	а	f	d	а	f
С	b	d			

Α	В	С
а	b	С
d	а	f
С	b	d
b	g	а

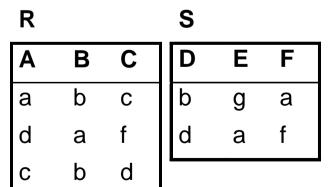
 $R \cup S$

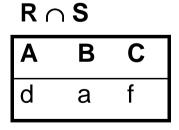
João Mendes Moreira



Intersecção

- 2 Intersecção $R \cap S$
 - o contém os tuplos que pertencem a R e a S simultaneamente
 - * R e S da mesma aridade
 - **▼** R e S com domínios compatíveis
 - $\circ R \cap S = R (R-S)$







Diferença

3 - Diferença

$$R - S$$

- o tuplos de R que não estão em S
 - * R e S da mesma aridade
 - **▼** R e S com domínios compatíveis

R

Α	В	С
а	b	С
d	а	f
С	b	d

S

D	E	F
b	g	а
d	a	f

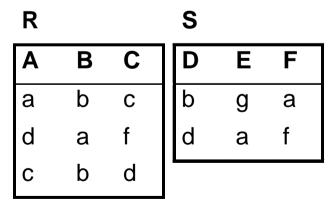
R-S

Α	В	С
а	b	С
С	b	d



Produto cartesiano

- 4 Produto cartesiano R×S
 - o aridades de R e S são k_1 e k_2 \Rightarrow aridade de R×S é k_1 + k_2



<u>KX</u>	<u> </u>				
Α	В	С	D	Ε	F
а	b	С	b	g	а
а	b	С	d	a	f
d	a	f	b	g	а
d	a	f	d	a	f
С	b	d	b	g	а
С	b	d	d	а	f



Projecção

5 - Projecção $\Pi_{i_1, i_2, ..., i_m}(R)$

- o para cada tuplo em R existe um tuplo na projecção com os componentes (e pela ordem) indicados pelos i_i
 - ullet se aridade de R for k então os $i_j \in 1, ...,$ k são distintos; aridade da projecção é m
 - × números de componentes podem ser substituídos por atributos, se existirem: $\Pi_{1,3}(R) = \Pi_{A,C}(R)$

K		
Α	В	С
а	b	С
d	а	f
С	b	d

$\Pi_{1,3}$	(R)
Α	С
а	С
d	f
С	d

Projecção escolhe colunas da tabela



Selecção

6 - Selecção $\sigma_{\rm F}({\rm R})$

- o contém os tuplos de R que satisfazem F
 - ▼ a fórmula F pode envolver
 - ⇒ operandos constantes ou número de componente (\$i)
 - \Rightarrow operadores aritméticos de comparação (<, =, >, \leq , \neq , \geq)
 - \Rightarrow operadores lógicos (\land , \lor , \neg) (e, ou, não)
 - × Números de componentes podem ser substituídos por atributos, se existirem: $\sigma_{\$2=b}(R) = \sigma_{B=b}(R)$ R

Α	В	С
а	b	С
d	а	f
С	b	d

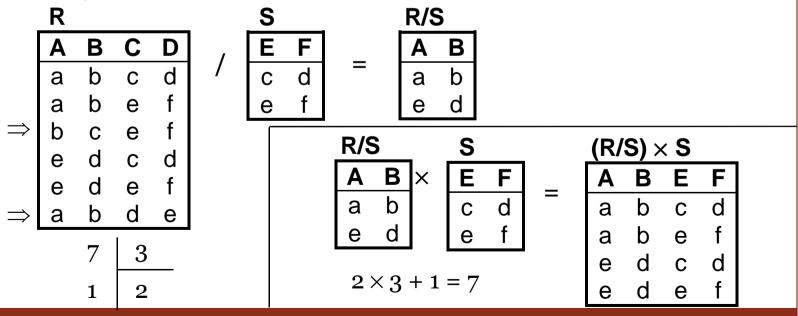
$\sigma_{$2=b}(R)$				
Α	В	С		
а	b	С		
С	b	d		

Selecção escolhe linhas da tabela



Quociente

- 7 Quociente R/S
 - × aridade de R é r e de S é s, r>s, S≠∅
 - o contém os (r-s)-tuplos $(a_1, ..., a_{r-s})$ tais que, para **todos** os s-tuplos $(a_{r-s+1}, ..., a_r)$ em S, o tuplo $(a_1, ..., a_r)$ está em R

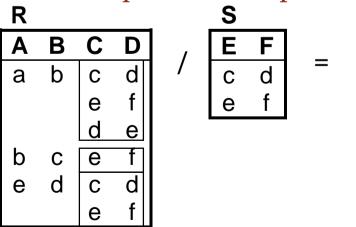




Quociente: explicação alternativa

- Forma de proceder à divisão
 - o reordenar as colunas de forma a que as últimas correspondam ao divisor
 - o ordenar a tabela pelas primeiras colunas
 - o cada subtuplo das primeiras colunas pertence ao resultado se o conjunto de subtuplos das últimas colunas que lhe corresponde

contiver o divisor



R/S

João Mendes Moreira FEUP



R

Α	В	C	U
a	b	С	d
a	b	е	f
b	С	е	f
е	d	С	d
е	d	е	f
а	b	d	е

S E F c d

A B a b b c e d

 TxS

 A
 B
 E
 F

 a
 b
 c
 d

 a
 b
 c
 d

 b
 c
 c
 d

 b
 c
 e
 f

 e
 d
 c
 d

 e
 d
 c
 d

 e
 d
 c
 f

Quociente: expressão

- \circ $T = \prod_{1, ..., r-s} (R)$ = universo dos tuplos possíveis no resultado
- \circ W = (T×S) R = todas as linhas T combinadas com S mas que não estão em R, i.e., em que a condição falha
- $\circ V = \Pi_{1, \dots, r-s} (W)$ = tuplos que não interessam
- \circ R / S = T V

= tuplos que interessam

reunindo numa só expressão algébrica

$$R / S = \Pi_{1, ..., r-s} (R) - \Pi_{1, ..., r-s} [(\Pi_{1, ..., r-s} (R) \times S) - R]$$

W A B E F b c c d

A B b c

R/S
A B
a b
e d



Exercícios

Quais as frases verdadeiras?

```
a - (R-S) ∪ S = R

b - (R-S) ∪ S ⊇ R

c - (R-S) ∪ (R∩S) = R

d - (R-S) ∪ (S-R) = (R∪S) - (R∩S)

e - (R/S) × S = R

f - (R/S) × S ⊆ R
```

Respostas

```
o Erradas - a, e
```

o Correctas - b, c, d, f



- x se a aridade de R for r e a de S s, a aridade da θ-junção é r+s
- o contém os tuplos do produto cartesiano de R por S tais que o componente i está na relação θ com o componente r+j (i.e., o correspondente ao j em S).
- Expressão da θ-junção

$$\circ R \bowtie_{i \in S} S = \sigma_{(r+j)} (R \times S)$$

- $R \bowtie_{i\theta j} S = \sigma_{\$i \theta \$(r+j)} (R \times S)$ × se θ for =, a operação designa-se equijunção
 - × (7, 8, 9) é um tuplo pendente de R pois não aparece na θ-junção

R			_	S	
Α	В	С		D	Е
1	2	3		3	1
4	5	6		6	2
7	8	9			

$ \begin{array}{ccc} R \bowtie S &= R \bowtie S \\ 2 &< 1 & B < D \end{array} $					
Α	В	С	D	Ε	
1	2	3	3	1	
1	2	3	6	2	
4	5	6	6	2	

 $P \bowtie Q - P \bowtie Q$



Junção natural

9 - Junção natural R ⋈ S

- x só é aplicável se os componentes dos tuplos em R e S forem designados por atributos.
- ▼ a operação implícita na junção natural é a igualdade dos atributos com o mesmo nome.
- cada par de atributos iguais dá origem a um único atributo, com o mesmo nome, no resultado
- o expressão:

$$R\bowtie S = \prod_{i_1, \dots, i_m} (\sigma_{R.A_1 = S.A_1.\wedge \dots \wedge .R.A_k = S.A_k} (R \times S))$$

x k é o número de atributos comuns a R (aridade r) e S (aridade s) e m= r+s-k

R		
Α	В	С
а	b	С
d	b	С
b	b	f
С	а	d

S		
В	С	D
b	С	d
b	С	е
а	d	b

 $R\bowtie S$

Α	В	С	D
а	b	С	d
а	b	С	е
d	b	С	d
d	b	С	е
С	а	d	b



Junção externa

- \circ tuplos pendentes, isto é desemparelhados, quer em R quer em S, desaparecem na θ-junção e na junção natural
- o a **junção externa (**θ- ou natural) inclui os tuplos pendentes de R ou S completados a nulos
 - × (7, 8, 9) é um tuplo pendente de R pois não aparece na θ-junção
 - ➤ (b, b, f) idem, na junção natural

R		
Α	В	С
1	2	3
4	5	6
7	8	9

S D E3 1
6 2

R I	\bowtie S	•		
В	<d< th=""><th></th><th></th><th></th></d<>			
A	В	C	D	Ε
1	2	3	3	1
1	2	3	6	2
4	5	6	6	2
7	8	9	Τ	\perp

K		
Α	В	С
а	b	С
d	b	С
b	b	f
С	а	d

	S		
	В	С	D
	b	С	d
	b	С	е
	а	d	b
,		•	

R	÷ 5	3	
Α	В	С	D
а	b	С	d
а	b	С	е
d	b	С	d
d	b	С	е
С	а	d	b
b	b	f	\perp



Semi-junção

- 10 Semi-junção R×S
 - o projecção nos atributos de R da junção natural de R e S
 - o $R \ltimes S = \prod_{R} (R \bowtie S)$
 - \times R em Π_R representa os atributos de R (o seu esquema); em R \bowtie S R representa a relação (a instância)
 - o outra expressão: $R \ltimes S = R \bowtie \Pi_{R \cap S}(S)$
 - o dá os tuplos de R que têm par em S

R		_	S		
Α	В	С	В	С	D
а	b	С	b	С	d
d	b	С	b	С	е
b	b	f	а	d	b
С	а	d			

R⋉S					
Α	В	С			
а	b	С			
d	b	С			
С	а	d			



Relações com atributos

- o na junção natural e na semi-junção os atributos são importantes; para os tornar explícitos escreve-se $R(A_1, ..., A_n)$
- o é possível renomear colunas e fazer junções naturais como:

$$S = S(B,C,D)$$

$$B \quad C \quad D$$

$$b \quad c \quad d$$

Ε	F	G	Н	I
а	d	b	С	d
а	d	b	С	е

- o uma junção natural entre duas relações sem atributos comuns redunda num produto cartesiano porque, após este, não há nenhuma selecção a fazer (equivalente a fazer uma selecção com a condição *True*)
- o R(A,B,C) \bowtie S(G,H,I) = R \times S

Leis algébricas

Reunião

o associativa: $R \cup (S \cup T) = (R \cup S) \cup T$ o comutativa: $R \cup S = S \cup R$

Produto cartesiano

o associativo: $R \times (S \times T) = (R \times S) \times T$

o não comutativo: $R \times \dot{S} \neq S \times \dot{R}$

Junção natural

- o associativa e comutativa (independência da ordem das colunas devida aos atributos): $\dot{R} \bowtie S = S \bowtie R$
- o Por isso \bowtie generaliza facilmente: $R = R_1 \bowtie ... \bowtie R_n$
 - \star R contém os tuplos μ tais que, para $1 \le i \le n$, μ restringido aos atributos de R_i é um tuplo de R_i

• θ - junção

o não é comutativa mas é associativa (no caso de os índices serem válidos)

$$R \bowtie_{i \theta_1 j} (S \bowtie_{k \theta_2 l} T) = (R \bowtie_{i \theta_1 j} S) \bowtie_{(r+k) \theta_2 l} T$$

Esquema relacional de Cursos



- 1 Curso(<u>codcurso</u>, designacur)
- 2 Disciplina(coddis, sigla, designadis)
- **3** Turma(<u>codcurso</u> → Curso, ano, letra)
- 4 Professor(bip, nome, morada, telefone, habilitação, grupo)
- 5 Aluno(bia, nome, morada, telefone, data_nasc, [codcurso, ano, letra]→Turma)
- 6 Plano(codcurso→Curso, coddis→Disciplina)
- 7 Inscrito(<u>coddis</u>→<u>Disciplina</u>, <u>bia</u>→<u>Aluno</u>, resultado)
- 8 Lecciona(bip → Professor, coddis → Disciplina, [codcurso, ano, letra] → Turma)

Só as tabelas 4 – 7 são necessárias para os exercícios seguintes.



- Álgebra Relacional pode ser usada como linguagem de interrogação à BD
- P1 Relativamente à BD "Cursos" (ver atrás),
 quais os nomes dos professores do 12º grupo?

```
\Pi_{\text{nome}} \ [\ \sigma_{\text{grupo} = '12'} \ (\ Professor\ )\ ]
```

• P2 - Quais os nomes e datas de nascimento dos alunos do curso 'CG1' nascidos antes de 1983?

```
\Pi_{\text{nome, data\_nasc}} [ \sigma_{\text{codcurso = 'CG1'} \land \text{data\_nasc < 1983-o1-o1}} ( Aluno ) ]
```

João Mendes Moreira FEUP



Perguntas com junção

- P3 Nomes dos alunos inscritos à disciplina 327?
 - o nenhuma relação contém nomes de alunos e códigos de disciplina
 - o mas a junção Aluno ⋈ Inscrito = R contém:
 - o R (<u>bia</u>, nome, morada, telefone, data_nasc, codcurso, ano, letra, coddis, resultado)

```
(i) - \Pi_{\text{nome}} [ \sigma_{\text{coddis} = 327} (Aluno \bowtie Inscrito )]
(ii) - \Pi_{\text{nome}} [ Aluno \bowtie \sigma_{\text{coddis} = 327} (Inscrito )]
```

o esta maneira de ligar informações no modelo de dados dá muita liberdade para exprimir perguntas arbitrárias mas exige uma fase de optimização para executar (ii) mesmo que a pergunta seja (i)

Núcleo da álgebra relacional : Π , σ , \bowtie

Extensões à álgebra relacional



Eliminação de duplicados

- $\circ R$
- $\circ R' = \delta(R)$
 - ▼ elimina os tuplos repetidos de R
- *P4 Quais os alunos inscritos a alguma disciplina?*
 - $\circ \Pi_{\text{nome}}$ [Aluno $\bowtie \delta$ [Π_{bia} (Inscrito)]]
 - A relação Inscrito contém os bi dos alunos tantas vezes quantas as cadeiras a que o aluno está inscrito
 - Assim, é necessário eliminar os bi repetidos para que os nomes dos alunos não apareçam repetidos

Extensões à álgebra relacional



Renomeações de atributos

- o R (A, B, C)
- o R'(X,Y,Z) = $\Pi_{X=A,Y=B,Z=C}$ [R(A,B,C)]
 - ➤ onde não houver ambiguidades, a simples menção dos atributos, em conjunto com o nome da relação, faz a renomeação

OU

- o R' (X, Y, Z) = R (A, B, C)OU
- o R' = $\Pi_{X = A, Y = B, Z = C}$ (R)
- expressões aritméticas (+, -, *, /)
- o este mecanismo serve para dar nomes a expressões

$$\circ S = \prod_{W = A * B - C, U = C/B, A} (R)$$
OU

$$\circ$$
 S (W, U, A)= $\Pi_{A * B - C, C/B, A}$ (R)



Expressões artméticas

• P5 – Obtenha a relação das inscrições com as classificações inflaccionadas de 20%.

o $\Pi_{\text{coddis, bia, resultado, novo = resultado*}1.2}$ (Inscrito)

Outro exemplo:

o $\Pi_{\text{coddis, bia, resultado, novo = resultado + (20-resultado)/10}}$ (Inscrito)

- nos parâmetros da projecção:
 - × no membro direito só podem ser usados nomes de atributos do argumento de Π; no esquerdo só pode estar um atributo (novo...)

João Mendes Moreira

Extensões à álgebra relacional



Agregações

- Operadores de agregação
 - o CNT (contagem), SUM (adição), AVG (média), MAX (máximo), MIN (mínimo)
- $S = \prod_{V = CNT(B)} (R)$
 - S(V) tem um único valor, o número de tuplos de R com valor não nulo no atributo B (CNT (*) conta todas as linhas) → toda a relação agregada
- $T = \Pi_{A, M = MAX(B)}(R)$
 - o T (A, M) tem tantos pares quantos os valores diferentes de A, sendo indicado para cada A o respectivo valor máximo de B (não nulo ...);
 - o é feita uma partição segundo os atributos de projecção sem operadores de agregação e cada classe é agregada num só tuplo;
 - o é possível misturar agregações e aritmética



Perguntas com agregação

- P6 Qual o número de inscrições, nota média das inscrições, soma de todas as notas e número de resultados não nulos ?
 - ο R(NI, M, T, NR) = $\Pi_{\text{CNT(*), AVG(resultado), SUM(resultado), CNT(resultado)}} \text{(Inscrito)}$
 - pode ser M ≠ T/NI se houver inscrições ainda sem resultado (valor nulo); tem que ser M = T/NR

Inscrito

coddis	bia	resultado
PA	97	14
PA	38	12
ITI	97	17
ITI	25	14
Н	97	10
Н	25	

R

NI	M	T	NR
6	13.4	67	5

João Mendes Moreira FEUP



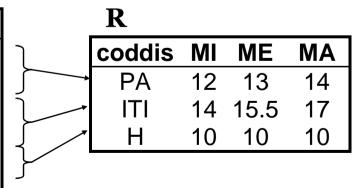
Agregação com partição

• P7 - Quais as notas mínima, média e máxima de cada disciplina (independentemente do aluno)?

o $R = \prod_{\text{coddis, MI = MIN(resultado), ME = AVG(resultado), MA = MAX(resultado)}}$ (Inscrito)

Inscrito

coddis	bia	resultado
PA	97	14
PA	38	12
ITI	97	17
ITI	25	14
H	97	10
Н	25	



João Mendes Moreira



Quantificação existencial

• P8 - Obtenha os códigos dos alunos com inscrição a **pelo menos uma** das disciplinas do curso LEEC.

o $\delta(\Pi_{BIA}(INSCRITO \bowtie \sigma_{CODCURSO= 'LEEC'}(PLANO)))$

João Mendes Moreira FEUP



Quantificação universal

- P9 Obtenha o código dos alunos com inscrição a **todas** as disciplinas do curso LEEC.
 - \circ A = Π_{BIA} (ALUNO)
 - conjunto dos alunos
 - o $D = \Pi_{CODDIS} [\sigma_{CODCURSO= 'LEEC'} (PLANO)]$
 - conjunto das disciplinas do curso LEEC
 - o A×D
 - conjunto de todos os pares (aluno, disciplina da LEEC)
 - o NI = $A \times D \Pi_{BIA, CODDIS}$ (INSCRITO)
 - pares (aluno, disciplina da LEEC) tais que o aluno não tem inscrição à disciplina
 - \circ R = A Π_{BIA} (NI)
 - resultado (notar a dupla subtracção)
- $R = \Pi_{BIA, CODDIS}(INSCRITO) / D$

Query language



More aggregation

- P10 Qual o nome e média actual do aluno com melhor média do curso 'LEEC'?
- Só a média actual dos alunos
 - o MA = $\Pi_{\text{bia, media = AVG(resultado)}}$ (Inscrito $\bowtie \sigma_{\text{CODCURSO= 'LEEC'}}$ (PLANO))
- Só a média máxima
 - \circ MaxM = $\Pi_{\text{media} = \text{MAX(media)}}$ (MA)
- Aluno cuja média é igual à média máxima
 - \circ R= $\Pi_{\text{nome, MA,media}}$ (Aluno \bowtie MA \bowtie MaxM)