
Álgebra Relacional

Sistemas de Informação I

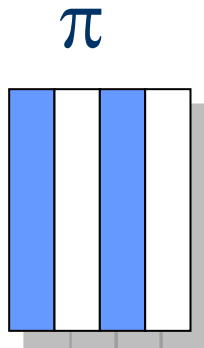
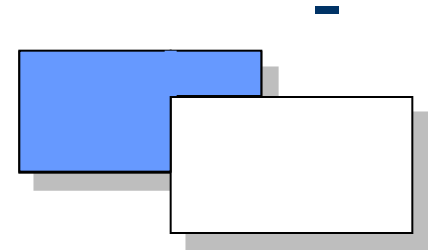
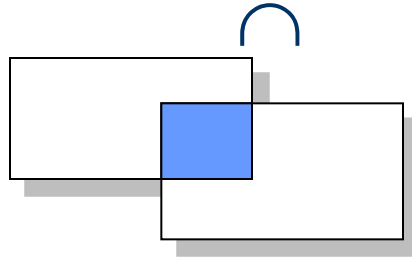
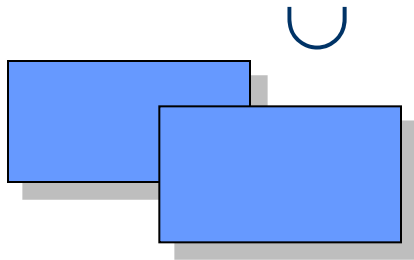
Álgebra Relacional

- Foi proposta por E. F. Codd
- É uma colecção de operações formais aplicáveis sobre Relações
- Como resultado da aplicação dessas operações obtêm-se sempre uma nova Relação
- Algumas dessas operações derivam directamente da teoria dos conjuntos
- Outras são específicas para o modelo relacional

Operações Relacionais

- Existentes no formalismo matemático da teoria dos conjuntos e aplicáveis pois uma Relação é um **conjunto** de tuplos:
 - União (operador \cup)
 - Intersecção (operador \cap)
 - Diferença (operador $-$)
 - Produto cartesiano (operador \times)
- Operações que foram elaboradas para o modelo relacional:
 - Selecção, restrição (operador σ)
 - Projecção (operador π)
 - Junção (operador \bowtie)
 - Divisão (operador \div)

Operações Relacionais (cont.)



\times

11	aa
22	bb
33	cc

$=$

11	aa
11	bb
11	cc
22	aa
22	bb
22	cc
33	aa
33	bb
33	cc

\div

A	B	C
a	t	y
b	t	y
c	x	z
d	t	y
e	x	z
b	x	z

$=$

A
b

Operações Relacionais (cont.)

- Formalmente, os nomes dos atributos das Relações operandos não são tidos em conta
- Cada atributo identifica-se pelo índice que ocupa na relação
- Na prática, e por questões de simplicidade e legibilidade, adopta-se o nome dos atributos para identificação
- Quando tal for ambíguo, utiliza-se o índice para identificação, por vezes entre parêntesis rectos

A	B	C
a	b	c
d	a	f
c	b	d

O atributo B, no contexto da relação correspondente, pode ser identificado por :

- B
- 2
- [2]

Seleccção ou restrição

- A operação de selecção “escolhe” um conjunto de tuplos de uma relação que verificam uma determinada condição
- Sendo r uma relação, esta operação representa-se da seguinte forma:
 - $\sigma_{\langle \text{condição booleana} \rangle}(r)$
- Exemplo: Quais as pessoas com idade superior a 25 anos?
 - $\sigma_{\text{idade} > 25}(\text{PESSOA})$

PESSOA

BI	Nome	Idade
123	José	25
456	José	70
789	Joana	15
909	Pedro	30

$\sigma_{\text{idade} > 25}(\text{PESSOA})$

BI	Nome	Idade
456	José	70
909	Pedro	30

Seleccção ou restrição (cont.)

- Cada condição booleana consiste numa sequências de cláusulas da forma:
 - Atributo op valor, em que valor pertence ao domínio do atributo
 - Atributo op Atributo
- As operações op podem ser:
 - $<, >, =, \geq, \leq, \neq$
- As cláusulas estão ligadas entre si por operadores lógicos:
 - \wedge (e), \vee (ou), \neg (negação)
- Quais as pessoas com idade superior a 25 anos e com nome José?
 - $\sigma_{\text{idade} > 25 \wedge \text{nome} = \text{José}} (\text{PESSOA})$

BI	Nome	Idade
456	José	70

Seleccção ou restrição – características

- A operação de selecção é uma operação unária, que tem como operando uma única relação
- O grau da relação resultante da aplicação da operação é igual ao grau da relação operando
- Uma sequência de operações σ pode ser escrita numa só que tem como condição a conjunção das condições das várias operações σ
 - $\sigma_{\text{cond1}} (\sigma_{\text{cond2}} (\dots (\sigma_{\text{condN}}(r)))) = \sigma_{(\text{cond1} \wedge \text{cond2} \wedge \text{condN})} (r)$
- A composição de operações σ é comutativa, ou seja:
 - $\sigma_{\text{cond1}} (\sigma_{\text{cond2}} (r)) = \sigma_{\text{cond2}} (\sigma_{\text{cond1}} (r))$

Quais as pessoas com idade superior a 25 anos e com nome José?

$$\sigma_{\text{idade} > 25} (\sigma_{\text{nome} = \text{José}} (\text{PESSOA})) = \sigma_{\text{idade} > 25 \wedge \text{nome} = \text{José}} (\text{PESSOA})$$

$$\sigma_{\text{idade} > 25} (\sigma_{\text{nome} = \text{José}} (\text{PESSOA})) = \sigma_{\text{nome} = \text{José}} (\sigma_{\text{idade} > 25} (\text{PESSOA}))$$

Projecção

- Operação de projecção escolhe um determinado subconjunto de atributos da relação operando
- Sendo r uma relação, esta operação representa-se da seguinte forma:
 - $\pi_{\langle \text{lista de atributos} \rangle}(r)$
- Quais os nomes e idades das pessoas?
 - $\pi_{\text{nome, idade}}(\text{PESSOA})$

PESSOA

BI	Nome	Idade
123	José	25
456	José	70
789	Joana	15
909	Pedro	30

$\pi_{\text{nome, idade}}(\text{PESSOA})$

Nome	Idade
José	25
José	70
Joana	15
Pedro	30

Projecção – características

- Quando no resultado de uma operação de projecção existirem vários tuplos iguais, apenas um é considerado:
 - Este procedimento é conhecido como “Eliminação de duplicados”
 - Com este procedimento é garantido que o resultado é uma relação, ou seja, um conjunto
- Se na lista de atributos da projecção estiver contida a chave da relação, a cardinalidade do resultado é a igual à cardinalidade da relação operando

Na composição de várias projecções

- Se lista1 está contida em lista2
 - $\pi_{\text{lista1}} (\pi_{\text{lista2}} (\text{PESSOA})) = \pi_{\text{lista1}} (\text{PESSOA})$
- Se lista1 não está contida em lista2
 - $\pi_{\text{lista1}} (\pi_{\text{lista2}} (\text{PESSOA})) = \text{operação inválida}$

Encadeamento de operações

- Escrever as operações numa única expressão
- As operações são aplicadas uma de cada vez a resultados intermédios

Para o exemplo anterior

Quais os nomes e idades das pessoas cuja idade é superior a 25 anos?

$\pi_{\text{nome, idade}} (\sigma_{\text{idade} > 25} (\text{PESSOA}))$

$\text{Raux} = \sigma_{\text{idade} > 25} (\text{PESSOA}) ; \pi_{\text{nome, idade}} (\text{Raux})$

BI	Nome	Idade
123	José	25
456	José	70
789	Joana	15
909	Pedro	30

$\pi_{\text{nome, idade}} (\sigma_{\text{idade} > 25} (\text{PESSOA}))$

Nome	Idade
José	70
Pedro	30

Álgebra Relacional – Relações exemplo

Considere os seguintes Esquemas de Relação e as respectivas Relações

- ALUNO (numeroAluno, nomeAluno, dataNascimentoAluno)
- DOCENTE (numeroDocente, nomeDocente, dataNascimentoDocente)

ALUNO

numeroAluno	nomeAluno	dataNascAluno
12345	Antonio	10-10-1970
43321	Nuno	21-09-1970
12231	Maria	12-05-1976

DOCENTE

numeroDocente	nomeDocente	dataNascDocente
12345	Antonio	10-10-1970
11223	Felizberta	21-09-1961

União

- A operação de união de duas Relações R1 e R2 tem como resultado uma Relação R, contendo os tuplos de R1 e R2
- Os tuplos repetidos são removidos
- Esta operação só é válida se R1 e R2 tiverem Esquemas compatíveis
- É uma operação comutativa: $R1 \cup R2 = R2 \cup R1$
- Formalmente escreve-se $R1 \cup R2 = \{t: t \in R1 \vee t \in R2\}$
- Quais o conjunto de todos os alunos e docentes?
 - $ALUNO \cup DOCENTE$

12345	Antonio	10-10-1970
43321	Nuno	21-09-1970
12231	Maria	12-05-1976
11223	Felizberta	21-09-1961

Diferença

- A operação de diferença entre duas Relações R1 e R2 tem como resultado uma Relação R, contendo os tuplos de R1 que não constam de R2
- Esta operação só é válida se R1 e R2 tiverem Esquemas compatíveis
- Formalmente escreve-se $R1 - R2 = \{t: t \in R1 \wedge t \notin R2\}$
- Qual o conjunto de todos os alunos que não são docentes?
 - ALUNO - DOCENTE

43321	Nuno	21-09-1970
12231	Maria	12-05-1976

Diferença - características

- Ao contrário da operação união, a operação diferença não é comutativa: $R1 - R2 \neq R2 - R1$
- Para o exemplo:

ALUNO - DOCENTE tem como resultado todos os alunos que não são docentes

43321	Nuno	21-09-1970
12231	Maria	12-05-1976

DOCENTE -ALUNO tem como resultado todos os docentes que não são alunos

11223	Felizberta	21-09-1961

Intersecção

- A operação intersecção entre duas Relações R1 e R2 tem como resultado uma Relação R constituída pelos tuplos comuns a R1 e R2
- Esta operação só é válida se R1 E R2 tiverem **esquemas compatíveis**
- Formalmente escreve-se $R1 \cap R2 = \{t: t \in R1 \wedge t \in R2\}$
- A intersecção pode ser feita recorrendo à diferença:
 - $R1 - (R1 - R2)$

Qual o conjunto de todos os alunos que também são docentes?

–ALUNO \cap DOCENTE

12345	Antonio	10-10-1970

União, diferença, intersecção: características comuns

- São operações binárias, ou seja, são operações efectuadas sobre duas Relações
- Os atributos da Relação resultante podem ser identificados pelo seu índice
- Para que as operações possam ser aplicadas, as Relações operando têm de ter Esquemas compatíveis
- Os Esquemas são compatíveis quando:
 - Têm o mesmo grau
 - Os atributos têm o mesmo domínio
- O Esquema da Relação obtida é igual ao das relações operando

Álgebra Relacional – Relações exemplo II

Considere os seguinte Esquemas de Relação e as respectivas Relações

- BANDA (codigo, nome, anoFormacao, genero)
- GENERO (codigo, designacao)

BANDA

Codigo	Nome	AnoFormacao	Genero
1	Metallica	1981	1
2	Madredeus	1991	2
3	Iron Maiden	1976	1
4	The Platters	1953	3

GENERO

Codigo	Designacao
1	Metal
2	Fado
3	Rock

Produto cartesiano

- A operação produto cartesiano tem como argumentos duas Relações R1 e R2
- Como resultado , obtém-se uma Relação R constituída pelas combinações possíveis dos tuplos de R1 com R2 (por esta ordem)
- Formalmente escreve-se
 - $R1 \times R2 = \{t1, t2: t1 \in R1 \wedge t2 \in R2\}$
- Se o grau de R1 for G1 e o de R2 for G2, então
 - O grau da Relação resultante de $R1 \times R2$ será igual a $G1 + G2$
- Se a cardinalidade de R1 for C1 e a de R2 for C2, então
 - A cardinalidade da Relação resultante de $R1 \times R2$ será igual a $C1 * C2$

Produto Cartesiano

Para as Relações BANDA e GENERO pretende-se saber qual a Relação resultante do produto cartesiano entre as duas?

Codigo	Nome	AnoFormacao	Genero	Codigo	Designacao
1	Metallica	1981	1	1	Metal
1	Metallica	1981	1	2	Fado
1	Metallica	1981	1	3	Rock
2	Madredeus	1991	2	1	Metal
2	Madredeus	1991	2	2	Fado
2	Madredeus	1991	2	3	Rock
3	Iron Maiden	1976	1	1	Metal
3	Iron Maiden	1976	1	2	Fado
3	Iron Maiden	1976	1	3	Rock
4	The Platters	1953	3	1	Metal
4	The Platters	1953	3	2	Fado
4	The Platters	1953	3	3	Rock

Produto cartesiano

- A aplicação de uma selecção à relação resultante de um produto cartesiano tem um grande interesse prático
- Esta composição visa restringir o resultado de uma combinação de duas tabelas
- Para o exemplo anterior pretende-se obter uma Relação que contenha a informação completa de uma BANDA e o seu respectivo GENERO: $\sigma_{\text{genero=codigo}}$ (**BANDA x GENERO**)

Codigo	Nome	AnoFormacao	Genero	Codigo	Designacao
1	Metallica	1981	1	1	Metal
2	Madredeus	1991	2	2	Fado
3	Iron Maiden	1976	1	1	Metal
4	The Platters	1953	3	3	Rock

Esta combinação de operações está na origem de uma outra, denominada de **Junção**

Junção

- A operação de junção entre duas Relações R1 e R2 é definida da seguinte forma: $R1 \bowtie_{\langle \text{condição junção} \rangle} R2$
- A condição de junção pode ser constituída pela conjunção de várias condições, cada uma da forma $i \theta j$, sendo i o i -ésimo atributo de R1 e j o j -ésimo de R2
- i e j têm o mesmo domínio
- O operador θ pode ser qualquer um dos seguintes operadores de comparação: $<, >, =, \geq, \leq, \neq$
- A operação junção definida desta forma designa-se de Junção Teta

Junção

- A operação de junção têm uma analogia directa com o encadeamento de duas operações (uma Seleção sobre um Produto cartesiano):
 - $R1 \bowtie_{i \theta j} R2 = \sigma_{A_i \theta A(N+j)} (R1 \times R2)$, onde N é o grau de R1
- Os tuplos que tiverem o valor NULL nos atributos de junção (i e j) não são contabilizados, ou seja, não pertencem à Relação resultante
- Do exemplo anterior
 - $BANDA \bowtie_{4=1} GENERO$

Codigo	Nome	AnoFormacao	Genero	Codigo	Designacao
1	Metallica	1981	1	1	Metal
2	Madredeus	1991	2	2	Fado
3	Iron Maiden	1976	1	1	Metal
4	The Platters	1953	3	3	Rock

Junção natural

- Considere-se os seguintes Esquemas de Relação e as respectivas Relações:
 - ALUNO (numeroAluno, nome, numeroMatricula)
 - AUTOMOVEL (numeroMatricula, modelo, cor)

Aluno

NumeroAluno	Nome	NumeroMatricula
12345	Felisberto	12-45-AA
18904	João	NULL
15444	Maria	00-45-EB
14566	Humberto	33-56-ST

Carro

NumeroMatricula	Modelo	Cor
12-45-AA	Honda Civic	Preto
33-56-ST	BMW Z3	Azul
94-11-FT	Ford Escort	Branco
00-45-EB	Mercedes C	Cinza

Junção natural (cont.)

- Para as relações anteriores pretende-se determinar os automóveis de cada aluno:
 - ALUNO $\bowtie_{3=1}$ CARRO

NumeroAluno	Nome	NumeroMatricula	Modelo	Cor
12345	Felisberto	12-45-AA	Honda Civic	Preto
15444	Maria	00-45-EB	Mercedes C	Cinza
14566	Humberto	33-56-ST	BMW Z3	Azul

- Quando a operação de junção é a igualdade, designa-se de *equijoin*
- Quando num *equijoin* os atributos de junção tem o mesmo nome, pode indicar-se a operação de junção da seguinte forma:
 - ALUNO $\bowtie_{\text{NumeroMatricula}}$ CARRO

A esta operação designa-se de junção natural

Junção natural (cont.)

- Para duas Relações R1 e R2 que partilham um atributo **x** com o mesmo domínio, designa-se de Junção natural a Relação que se obtêm juntando todos os tuplos de R1 e R2 que têm o mesmo valor para esse atributo **x**
- Formalmente:
 - $\{ \langle t1.A1, \dots, t2.B1, \dots, t1.X \rangle : t1 \in R1 \wedge t2 \in R2 \wedge t1.X = t2.X \}$
- Quando se omite o atributo (ou atributos) sobre o qual irá ser efectuada a junção, serão considerados os atributos R1 e R2 que têm o mesmo nome e tipo

Deve ser tomado algum cuidado quando se omitem os atributos, pois embora por vezes com o mesmo nome e domínio, a semântica associada é diferente !

Semi-Junção

- A operação de Semi-Junção, ou Junção Parcial, resulta da projecção dos atributos de R1 sobre a Relação resultante de uma Junção Natural entre as Relações R1 e R2
- Formalmente
 - $R1 \bowtie_x R2 = \pi_{R1.A1, \dots, R1.An} (R \bowtie_x S)$
- Quais os alunos que têm pelo menos um automóvel:
 - $ALUNO \bowtie_{\text{NumeroMatricula}} CARRO$

NumeroAluno	Nome	NumeroMatricula
12345	Felisberto	12-45-AA
18904	João	NULL
15444	Maria	00-45-EB
14566	Humberto	33-56-ST



NumeroMatricula	Modelo	Cor
12-45-AA	Honda Civic	Preto
33-56-ST	BMW Z3	Azul
94-11-FT	Ford Escort	Branco
00-45-EB	Mercedes C	Cinza

NumeroAluno	Nome	NumeroMatricula
12345	Felisberto	12-45-AA
15444	Maria	00-45-EB
14566	Humberto	33-56-ST

Junção Externa

- Nas operações de junção apresentadas, os tuplos de uma Relação R1 que não tenham associação noutra Relação R2 não constavam da Relação resultante

Será que este comportamento é sempre desejado?

- Para resolver esta particularidade foi criada uma operação de junção que aplicada a duas relações R1 e R2
- Quando os tuplos que tem associados são incluídos no resultados, como o resultado de uma junção
- Os outros são igualmente incluídos na Relação resultando, sendo completados com NULL nos atributos inexistentes
- Esta operação designa-se de Junção Externa:
 - $R1 \bowtie_{\theta} R2$

Junção Externa (cont.)

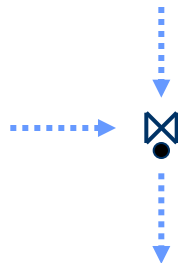
- Para as relações R1 e R2

R1

Codigo	Marca	Categoria
12345	Toyota Corolla	2
18904	Honda Accord	3
14566	Mercedes E3500	4

R2

Codigo	Descricao
1	1000CC
2	1500CC
3	2000CC
4	2500CC



Quais as marcas de automóveis e categorias existentes, e para cada automóvel as categorias superiores à sua?

$$R1 \bowtie_{[3] < [1]} R2$$

Codigo	Marca	Categoria	Codigo	Descricao
12345	Toyota	2	3	2000CC
12345	Toyota	2	4	2500CC
18904	Honda	3	4	2500CC
14566	Mercedes	4	NULL	NULL
NULL	NULL	NULL	1	1000CC
NULL	NULL	NULL	2	1500CC

Junção Externa à Esquerda e Direita

- Por vezes apenas os tuplos de uma das Relações se querem manter, numa Junção Externa
- Na operação de Junção à Esquerda (*Left Outer Join*) apenas os tuplos da Relação R1, à esquerda do operador, são totalmente considerados no resultado. Os tuplos de R2 aparecem quando verificam a condição

- $R1 \bowtie_{i \theta j} R2$

- Na operação de Junção à Direita (*Right Outer Join*) apenas os tuplos da Relação R2, à direita do operador, são totalmente considerados no resultado. Os tuplos de R1 aparecem quando verificam a condição

- $R1 \bowtie_{i \theta j} R2$

Junção Externa à Esquerda e Direita (cont.)

- Para as relações anteriores, pretende-se saber quais as categorias existentes e para essas quais os veículos existentes
- Ou seja
 - $R1 \bowtie_{3=1} R2$

Codigo	Marca	Categoria	Codigo	Descricao
12345	Toyota Corolla	2	2	1500CC
18904	Honda Accord	3	3	2000CC
14566	Mercedes E3500	4	4	2500CC
NULL	NULL	NULL	1	1000CC

Divisão

- A operação de divisão de duas Relações $R1$ e $R2$ tem como resultado uma Relação R , contendo um conjunto de tuplos que combinados com $R2$ dá um tuplo existente em $R1$
- A relação R é quociente de $R1 \div R2$
- Os tuplos de R são constituídos apenas pelos atributos de $R1$ que não fazem parte de $R2$
- Esta operação só é válida se $R1$ e $R2$ tiverem Esquemas compatíveis em divisão, ou seja, sendo X o conjunto de atributos de $R1$ e Y o conjunto de atributos de $R2$, essa compatibilidade existe quando $X \subseteq Y$
- Formalmente:
 - sendo $R1(A_1, \dots, A_p, A_{p+1}, A_n)$ e $R2(A_{p+1}, \dots, A_n)$ duas Relações, a operação $R1 \div R2$ tem como resultado uma Relação R tal que o tuplo que resulta de da concatenação de R com $R2$ dá um tuplo existente em $R1$

Divisão

- Para as seguintes relações R1 e R2:

R1

Fornecedor	Produto
1	Feijão
2	Arroz
3	Massa
1	Massa
2	Feijão
1	Arroz

R2

Produto
Feijão
Arroz
Massa

Qual o fornecedor que fornece todos os produtos?

○ $R1 \div R2$

Fornecedor
1

Operações agregadoras

- Para a seguinte relação R1

R1

Codigo	Nome	Ordenado	Departamento
12345	Felisberto	700	Contabilidade
12333	Maria	650	Secretariado
99912	João	700	Contabilidade
14566	Nuno	900	Secretariado
66277	Joana	1000	Tesouraria
22300	Manuel	700	Contabilidade

Como responder à questão:
Qual a média de ordenados ?

- É necessário utilizar funções especiais, designadas funções agregadoras, como

- **Count** – contagem de tuplos
- **Sum** – cálculo da soma
- **Average** – cálculo da média

- **Minimum** – identificação do mínimo
- **Maximum** – identificação do máximo

Operações agregadoras e operador de agregação

- Outra questão poderá ser o qual o maior dos ordenados, agrupados por departamento?
- A resposta implica agrupar um conjunto de tuplos e sobre um dos atributos desse tuplos aplicar uma operação agregadora
- A definição do operador \mathfrak{A} irá permitir a escrita destas questões
- A aplicação do operador faz-se sobre uma única Relação R e é representado da seguinte forma:
 - $\langle \text{Atributos para agrupamentos} \rangle \mathfrak{A} \langle \text{funções de agregação} \rangle (R)$
- A questão acima colocada pode ser escrita da forma
 - $\text{departamento} \mathfrak{A} \text{Maximum(ordenado)} (R1)$

Operações agregadoras

- <Atributos para agrupamento> consiste na lista de atributos sobre os quais se pretende agrupar os valores
- <Funções de agregação> consiste numa lista de pares (<função> <atributo>)
- Na relação resultante existem os atributos presentes em <Atributos para agrupamento> e os atributos que fazem parte dos pares (<função> <atributo>)

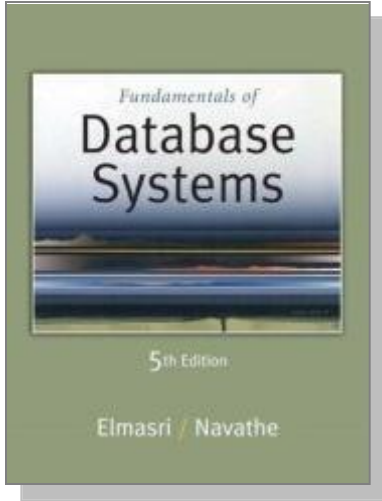
departamento \bowtie Maximum(ordenado) (R1)

Departamento	Ordenado
Contabilidade	700
Secretariado	900
Tesouraria	1000

\bowtie AVERAGE(ordenado) (R1)

775

Bibliografia



Fundamentals of Database System (5th Edition)

R. Elmasri, Shamkant B. Navathe

Addison Wesley, 2003