Banco de Dados

Álgebra Relacional

João Eduardo Ferreira Osvaldo Kotaro Takai Marcelo Finger

Tópicos

- Apresentação
- Seleção
- Projeção
- Renomeio de Atributos
- União, Subtração e operações de conjuntos
- Junção: Natural e outros tipos
- Divisão

Introdução

- A Álgebra Relacional: operações para consultas sob relações.
- As operações são divididas em dois grupos:
 - Operações da Teoria de Conjuntos
 - Operações desenvolvidas especificamente para Bancos de Dados Relacionais

MDR Companhia

ce





→ce

DNUM

- Utilizada para selecionar, segundo alguma condição, tuplas de uma relação.
- Exemplos:
 - Selecionar os empregados que trabalham para o departamento 4:
 - O doum = 4 (EMPREGADO)
 - Selecionar empregados que tenham salário maior que 3000
 - **O** _{SALÁRIO > 3000} (EMPREGADO)

- Para especificar as condições, podemos utilizar:
 - Valor Constante
 - Nome de Atributo
 - Os operadores relacionais: $\{=,<,\leq,\geq,\neq\}$
 - Os operadores lógicos: {AND, OR, NOT}

Exemplo:

Selecionar os empregados que trabalham no departamento 4 e ganham mais de 2500 ou aqueles que trabalham no departamento 5 e ganham mais que 3000.

PNOME	MNOME	SNOME	NSS	DATANASC	ENDEREÇO	SEXO	SALARIO	NSSSUPER	DNUM
John	В	Smith	12345678	09-JAN-55	R. A, 1	M	3000	333445555	5
Franklin	Т	Wong	33344555	08-DEZ-45	R. B, 2	M	4000	888665555	5
Alícia	J	Zelaya	99988777	19-JUL-58	Av. C, 3	F	2500	987654321	4
Jennifer	S	Wallace	98765432	20-JUN-31	Trav. D, 4	F	4300	888665555	4
Ramesh	K	Narayan	66688444	15-SET-52	R. E, 5	M	3800	333445555	5
Joyce	Α	English	45345345	31-JUL-62	R. F, 6	F	2500	333445555	5
Ahmad	V	Jabbar	98798798	29-MAR-59	Av G, 7	M	2500	987654321	4
James	E	Borg	88866555	10-NOV-27	Av H, 8	M	5500	null	3

Consulta:

(NDEP = 4 AND SALÁRIO > 2500) OR (NDEP = 5 AND SALÁRIO > 3000) (EMPREGADO)

Resultado:

PNOME	MNOME	SNOME	NSS	DATANASC END	EREÇO SEXO	SALARIO	NSSSUPER	NDEP
Franklin	т	Wong	33344555	08-DEZ-45 R. B,	2 M	4000	888665555	5
Jennifer	S	Wallace	98765432	20-JUN-31 Trav.	D, 4 F	4300	888665555	4
Ramesh	K	Narayan	66688444	15-SET-52 R. E,	5 M	3800	333445555	5

- Características e Propriedades do SELECT
 - É um operador unário. Seleciona tuplas de somente uma relação
 - O esquema da relação resultante é o mesmo da relação original
 - É comutativa: Pode-se trocar SELECT em cascata pela conjuntiva AND
 - $\mathbf{\sigma}$ <cond₂>($\mathbf{\sigma}$ <cond₂>(\mathbf{R})) = $\mathbf{\sigma}$ <cond₁>AND<cond₂>(\mathbf{R})

Operador PROJECT - π

- Enquanto o operador SELECT seleciona tuplas de uma relação, o operador PROJECT seleciona colunas de uma relação
- Por exemplo: Projetar os atributos SNOME, PNOME e SALÁRIO da relação EMPREGADO.

PNOME	MNOME	SNOME	NSS	DATANASC ENDEREÇO	SEXO	SALARIO	NSSSUPER	NDEP
John	В	Smith	123456789	09-JAN-55 R. A, 1	M	3000	333445555	5
Franklin	T	Wong	333445555	08-DEZ-45 R. B, 2	M	4000	888665555	5
Alícia	J	Zelaya	999887777	19-JUL-58 Av. C, 3	F	2500	987654321	4
Jennifer	S	Wallace	987654321	20-JUN-31 Trav. D, 4	F	4300	888665555	4
Ramesh	K	Narayan	666884444	15-SET-52 R. E, 5	M	3800	333445555	5
Joyce	Α	English	453453453	31-JUL-62 R. F, 6	F	2500	333445555	5
Ahmad	V	Jabbar	987987987	29-MAR-59 Av G, 7	M	2500	987654321	4
James	Е	Borg	888665555	10-NOV-27 Av H, 8	M	5500	null	3

Operador PROJECT - π

Consulta:

π PNOME, SNOME, SALÁRIO (EMPREGADO)

Resultado:

PNOME	SNOME	SALARIO
John	Smith	3000
Franklin	Wong	4000
Alícia	Zelaya	2500
Jennifer	Wallace	4300
Ramesh	Narayan	3800
Joyce	English	2500
Ahmad	Jabbar	2500
James	Borg	5500

Operador PROJECT - π

- O operador PROJECT remove quaisquer tuplas duplicadas da relação resultante
- Por exemplo: Projetar SEXO e SALÁRIO da relação EMPREGADO

 $\pi_{_{_{\mathsf{SEXO},\,\mathsf{SAL\acute{A}RIO}}}}(\mathsf{EMPREGADO})$

SEXO	SALARIO
M	3000
M	4000
F	2500
F	4300
M	3800
M	2500
M	5500

Operador PROJECT - π

- Características e propriedades do operador PROJECT:
 - O número de tuplas resultante sempre será igual ou menor que a quantidade de tuplas da relação original
 - Não é comutativa
 - Caso

$$\pi_{\text{clista1}}$$
 (π_{clista2} (R)) = π_{clista1} (R)

Sequência de Operações

- Podemos combinar os operadores em uma única expressão para realizar uma consulta
- Por exemplo: Recuperar o PNOME, SNOME e SALÁRIO de todos os empregados que trabalham no departamento 5.

PNOME	MNOME	SNOME	NSS	DATANASC	ENDEREÇO	SEXO	SALARIO	NSSSUPER	NDEP
John	В	Smith	123456789	09-JAN-55	R. A, 1	M	3000	333445555	5
Franklin	T	Wong	333445555	08-DEZ-45	R. B, 2	M	4000	888665555	5
Alícia	J	Zelaya	999887777	19-JUL-58	Av. C, 3	F	2500	987654321	4
Jennifer	S	Wallace	987654321	20-JUN-31	Trav. D, 4	F	4300	888665555	4
Ramesh	K	Narayan	666884444	15-SET-52	R. E, 5	M	3800	333445555	5
Joyce	Α	English	453453453	31-JUL-62	R. F, 6	F	2500	333445555	5
Ahmad	V	Jabbar	987987987	29-MAR-59	Av G, 7	M	2500	987654321	4
James	E	Borg	888665555	10-NOV-27	Av H, 8	M	5500	null	3

Sequência de Operações

Consulta:

$$\pi_{\text{PNOME, SNOME, SALÁRIO}}(\sigma_{\text{DNUM=5}}(\text{EMPREGADO}))$$

Resultado:

PNOME	SNOME	SALARIO
John	Smith	3000
Franklin	Wong	4000
Ramesh	Narayan	3800
Joyce	English	2500

Relações Intermediárias

Podemos criar relações intermediárias para explicitar a seqüência de operações:

$$\begin{aligned} & \mathsf{DEP5_EMPS} \leftarrow \sigma_{\mathsf{NDEP=5}}(\mathsf{EMPREGADO})) \\ & \mathsf{RESULT} \leftarrow \pi_{\;\;\mathsf{PNOME,\;\;SNOME,\;\;SAL\acute{A}RIO}}(\mathsf{DEP5_EMPS}) \end{aligned}$$

Renomeando Atributos

Ao criar uma nova relação, podemos renomear os atributos da relação resultante.

```
\overline{\mathsf{DEP5}\_\mathsf{EMPS}} \leftarrow \overline{\sigma_{\mathsf{NDEP=5}}}(\mathsf{EMPREGADO}))
```

 $\overline{\text{RESULT}(\text{NOME, SOBRENOME, SALÁRIO})} \leftarrow \overline{\pi}_{\text{PNOME, SNOME, SALÁRIO}}(\overline{\text{DEP5_EMPS}})$

RESULT

NOME	SOBRENOME	SALÁRIO
John	Smith	3000
Franklin	Wong	4000
Ramesh	Narayan	3800
Joyce	English	2500

- Os operadores da Teoria dos Conjuntos aplicam-se ao modelo relacional pois uma relação é como um conjunto de tuplas
- Por exemplo: Recuperar o NSS dos empregados que trabalham no departamento 5 ou, indiretamente supervisionem empregados que trabalham no departamento 5

```
DEP5_EMPS\leftarrow \sigmaNDEP=5 (EMPREGADO)

RESULT1\leftarrow \pi NSS (DEP5_EMPS)

RESULT2(NSS)\leftarrow \pi NSSSUPER (DEP5_EMPS)

RESULT\leftarrowRESULT1 \cup RESULT2
```

- Os operadores são:
 - R∪S União (todas as tuplas de R e todas de S)
 - R∩S Intersecção (todas as tuplas comuns a R e S)
 - R–S Diferença (todas as tuplas de R que não estão em S)
 - R x S Produto Cartesiano (combinação das tuplas de R com as de S)

Os operadores são: ROS União (todos as tu R e todas de S) $-R \cap S$ a ReS) UNIÃO R-S - Dife que não estão **INTERSECÇÃO** em 😜 **DIFERENÇA** ■ R x S – Pre Jao das tuplas PRECISAM QUE R E S SEJAM de R co COMPATÍVEIS NA UNIÃO

- As operações de união, intersecção e diferença são operações bastante intuitivas em suas aplicações.
- No entanto, o produto cartesiano não é tão intuitivo.

Para facilitar, considere o seguinte exemplo:

SNOME	<u>NSS</u>		<u>NSSEMP</u>	<u>NOMEDEPENDENT</u>
Smith	12345678	X	33344555	E lice
Wong	33344555	71	98765432	Abner
	5		12345678	Alice
			12345678	Elizabeth
			9	



SNOME	<u>NSS</u>	<u>NSSEMP</u>	<u>NOMEDEPENDENTE</u>
Smith	12345678	33344555	Alice
Smith	12345678	98765432	Abner
Smith	12345678	12345678	Alice
Smith	12345678	12345678	Elizabeth
Wong	33344555	33344555	Alice
Wong	33344555	98765432	Abner
Wong	33344555	12345678	Alice
Wong	33344555	12345678	Elizabeth
	5	9	

Recuperar, para cada empregado do sexo feminino, uma lista de nomes de seus dependentes

```
\begin{split} & \text{EMP\_FEM} \leftarrow \sigma_{\text{SEXO='F'}} \text{ (EMPREGADO)} \\ & \text{EMP\_NOMES} \leftarrow \pi_{\text{PNOME, SNOME, NSS}} \text{ (EMP\_FEM)} \\ & \text{EMP\_DEP} \leftarrow \text{EMP\_NOMES X DEPENDENTE} \\ & \text{DEP\_ATUAL} \leftarrow \sigma_{\text{NSS=ENSS}} \text{ (EMP\_DEP)} \\ & \text{RESULT} \leftarrow \pi_{\text{PNOME, SNOME, NOMEDEPENDENTE}} \text{ (DEP\_ATUAL)} \end{split}
```

Ver resultado na página 113 do livro do Elmasri & Navathe

Operador Join (►<)

- □ O *join* (▷<) é um dos operadores mais úteis da Álgebra Relacional.
- Normalmente o join é utilizado para combinar informações de duas ou mais relações.
- O join pode ser definido como um produto cartesiano seguido por uma seleção.

Operador Join

- Por exemplo, a consulta:
 - Recuperar os nomes de gerentes de cada departamento

```
DEP_EMP \leftarrow DEPARTAMENTO X EMPREGADO DEPT_GER \leftarrow \sigma_{\text{gernss=nss}} DEP_EMP RESULT \leftarrow \pi_{\text{pnome}} (DEPT_GER)
```

 $\mathsf{DEP_EMP} \leftarrow \mathsf{DEPARTAMENTO} \, \triangleright \, \triangleleft_{\mathsf{ERNSS=NSS}} \, \mathsf{EMPREGADO} \\ \mathsf{RESULT} \leftarrow \, \boldsymbol{\mathcal{T}}_{\mathsf{PNOME}}(\mathsf{DEPT_EMP})$

Operador Join

- Um outro exemplo:
 - Recuperar, para cada empregado do sexo feminino, uma lista de nomes de seus dependentes

```
 \begin{split} & \text{EMP\_FEM} \leftarrow \pmb{\sigma}_{\text{Sexo='F'}} \text{ (EMPREGADO)} \\ & \text{EMP\_NOMES} \leftarrow \pmb{\pi}_{\text{PNOME, SNOME, NSS}} \text{ (EMP\_FEM)} \\ & \text{EMP\_DEP} \leftarrow \text{EMP\_NOMES} \ \pmb{X} \ \text{DEPENDENTE} \\ & \text{DEP\_ATUAL} \leftarrow \pmb{\sigma}_{\text{NSS=ENSS}} \text{ (EMP\_DEP)} \\ & \text{RESULT} \leftarrow \pmb{\pi}_{\text{PNOME, SNOME, NOMEDEPENDENTE}} \text{ (DEP\_ATUAL)} \\ \end{aligned}
```

Operador Join

- Um outro exemplo:
 - Recuperar, para cada empregado do sexo feminino, uma lista de nomes de seus dependentes

$$\mathsf{EMP_FEM} \leftarrow \sigma_{\mathsf{SEXO}=\mathsf{'F'}} \, (\mathsf{EMPREGADO})$$

$$\mathsf{EMP_NOMES} \leftarrow \pi_{\mathsf{PNOME, SNOME, NSS}} \, (\mathsf{EMP_FEM})$$

$$\mathsf{RESULT} {\leftarrow} \pi_{\mathsf{PNOME, SNOME, NOMEDEPENDENTE}}(\mathsf{DEP_ATUAL})$$

Operador Equijoin

- É comum encontrar JOIN que tenham somente comparações de igualdade.
- Quando isso ocorre, o JOIN é chamado EQUIJOIN
- Note que no resultado de uma EQUIJOIN haverá, sempre, um ou mais pares de atributos com valores idênticos.

DEPT_GER ← DEPARTAMENTO ► ☐ GERNSS=NSS EMPREGADO

DNOME	DNUMERO	NSSGER	 PNOME	MNOME	SNOME	NSS	
Pesquisa	5	333445555	 Franklin	Т	Wong	333445555	
Administrativo	4	987654321	 Jennifer	S	Wallace	987654321	
Gerencial	1	888665555	 James	Е	Borg	888665555	

Operação NATURAL JOIN *

- Devido a tal duplicidade ser desnecessária, uma nova operação foi criada: NATURAL JOIN.
- O NATURAL JOIN (*), é um EQUIJOIN seguido da remoção de atributos desnecessários.
- A forma geral desse operador é:
 - Q←R * (lista1), (lista2) S, onde:
 - lista1 especifica os atributos de R e
 - lista2 os atributos de S.
- Na relação resultante, os atributos da *lista2* não irão aparecer.

Operação NATURAL JOIN *

- Pode-se continuar a especificar o sinal de igualdade na condição, apesar de ser desnecessária. Exemplo:
 - PROJ_DEPT ← PROJETO * DNUM = DNÚMERO DEPARTAMENTO OU
 - PROJ_DEPT ← PROJETO * (DNUM), (DNÚMERO) DEPARTAMENTO
- Pode-se omitir as listas de atributos.
 - Nesses casos, o operador irá considerar para a condição, os atributos que tiverem o mesmo nome em ambas as relações. Exemplo:
 - □ DEPT_LOCS ← DEPARTAMENTO * LOCAIS_DEPTO

- A operação de divisão é útil para um tipo especial de consulta que ocorre com freqüência. Por exemplo:
 - Recuperar os nomes de empregados que trabalham em todos os projetos em que John Smith trabalha.
- Primeiro, obtemos o conjunto de todos os projetos onde John Smith trabalha:
 - SMITH $\leftarrow \mathbf{O}_{PNOME='John'\ AND\ SNOME='Smith'}$ (EMPREGADO)
 - SMITH_PNO $\leftarrow \pi_{_{\mathsf{PNO}}}$ (TRABALHA_EM * $_{_{\mathsf{NSSEMP}\,=\,\mathsf{NSS}}}$ SMITH)

```
SMITH_PNO PNO 1 2
```

- Obter a relação com o NSS e PNO dos empregados
 - NSS_PNRO $\leftarrow \pi_{\text{PNO, NSSEMP}}$ (TRABALHA_EM)

NSS_PNRO	NSSEMP	PNO
	123456789	1
	123456789	2
	666884444	3
	453453453	1
	453453453	2
	333445555	2
	333445555	3
	333445555	10
	333445555	20
	999887777	30
	999887777	10
	987987987	10
	987987987	30
	987654321	30
	987654321	20
	8886655555	20

- Finalmente, aplicar a divisão:
 - NSS_DESEJADOS(NSS) ← NSS_PNRO ÷ SMITH_PNO

NSS_PNRO	NSSEMP	PNO
	123456789	1
	123456789	2
	666884444	3
	453453453	1
	453453453	2
	333445555	2
	333445555	3
	333445555	10
	333445555	20
	999887777	30
	999887777	10
	987987987	10
	987987987	30
	987654321	30
	987654321	20
	8886655555	20

SMITH_PNO	PNO	NSS_DESEJADO	NSS
	1		123456789
	2		453453453

- A divisão pode ser escrita em termos dos operadores básicos: π, x e –.
- Sejam duas instâncias de relação A(x, y) e B(y).
- O resultado de A÷B contém todos os valores x de A que não são desqualificados.
 - Um valor x é desqualificado se, ao anexar um valor y de B, resultar em tuplas <x, y> que não estão em A.
- A seguinte expressão permite fazer isso:
 - XDESQUALIFICADO $\leftarrow \pi_{x}((\pi_{x}(A)xB) A)$

- Agora, basta tirar de A as tuplas desqualificadas, resultando na divisão:
 - $A \div B \leftarrow \pi_{\times}(A) XDESQUALIFICADO$
- Assim
 - $A \div B \leftarrow \pi_{\times}(A) \pi_{\times}((\pi_{\times}(A) \times B) A)$

Funções de Agregação

- Funções agregadas recebem como entrada um conjunto de tuplas e retornam um único valor
 - COUNT
 - SUM
 - AVERAGE
 - MAXIMUM
 - MINIMUM
- □ Para especificar uma função agregada, utilizamos o operador FUNCTION - S

Funções de Agregação

- Por exemplo: Recuperar para cada departamento, o número de empregado e sua média salarial.
 - R(DNO, NRO_EMPS, MÉDIA) DNUM S COUNT NSS, AVERAGE SALÁRIO (EMPREGADO)



Atributo de Agrupamento

R	DNO	NRO_EMPS	MÉDIA
	5	4	3325
	4	3	3100
	1	1	5500

Funções de Agregação

Se nenhum atributo de agrupamento for especificado, as funções de agregação irão ser aplicadas para todas as tuplas da relação

$$R \leftarrow \Im$$
 count NSS, average Salário (EMPREGADO)

R	COUNT_NSS	AVERAGE_SALÁRIO
	8	3512.5

- A clausura recursiva é uma operação que não pode ser definida na álgebra relacional.
- Ela ocorre quando se tem relacionamentos recursivos. Por exemplo:
 - Recuperar todos os supervisionados de um empregado *e* em todos os níveis, isto é:
 - todos os empregados *e'* diretamente supervisionados por *e*, todos os empregados *e'* diretamente supervisionados por *e'*, e assim por diante.

- É simples especificar, na álgebra relacional, todos os empregados supervisionados por *e* num nível específico
- No entanto, não é possível especificar todos os supervisionados em todos os níveis

- Por exemplo:
 - Obter o NSS de todos os empregados diretamente supervisionados por James Borg
 - BORG_NSS $\leftarrow \pi_{\text{NSS}}$ ($\sigma_{\text{PNOME = 'James' AND SNOME = 'Borg'}}$ (EMPREGADO))
 - SUPERVISÃO(NSS1, NSS2) $\leftarrow \pi_{\text{NSS, NSSSUPER}}$ (EMPREGADO)
 - RESULT1 $\leftarrow \pi_{\text{NSS1}}$ (SUPERVISÃO $\triangleright \triangleleft_{\text{NSS2} = \text{NSS}}$ BORG_NSS)

- Para recuperar todos os supervisionados por Borg no nível 2, basta aplicar o JOIN ao resultado da primeira consulta:
 - RESULT2 $\leftarrow \pi_{\text{NSS1}}$ (SUPERVISÃO $\triangleright \triangleleft_{\text{NSS2} = \text{NSS}}$ RESULT1)
- Para obter todos os funcionários supervisionados por Borg nos níveis 1 e 2 basta aplicar a operação de união
 - RESULT3 ← RESULT1 ∪ RESULT2

Inner Joins (Junções Internas)

- Os operadores Join vistos até agora, os quais apenas tuplas que satisfazem a condição de junção são mantidas no resultado, são conhecidas como junções internas (*inner joins*).
- Por exemplo, no NATURAL JOIN, R * S, apenas a tuplas de R que correspondem as tuplas em S aparecem no resultado.

Outer Joins (Junções Externas)

- Junções externas podem ser utilizadas quando queremos manter todas as tuplas de R, S ou de ambas no resultado do Join, independentemente de existirem tuplas correspondentes na outra relação.
- Por exemplo, considere a consulta:
 - Obter a lista de nomes de todos os empregados e o nome dos departamentos que gerenciam. Se gerenciarem algum departamento. Se não gerenciarem nenhum departamento, indicar com um valor null.

Outer Joins (Junções Externas)

- Para a consulta, podemos utilizar a junção externa à esquerda (*Left Outer Join*), indicada por ▶ ◄ :
 - TEMP ← (EMPREGADO → ► NSS=GERNSS DEPARTAMENTO)
 - $\begin{tabular}{ll} \hline & RESULTADO \leftarrow \pi \\ & \hline \end{array}_{\tt PNOME, \, MNOME, \, SNOME, \, DNOME} (TEMP) \\ \hline \end{tabular}$

RESULTAD O	PNOME	MNOME	SNOME	DNOME
	John	В	Smith	Null
	Franklin	Т	Wong	Pesquisa
	Alícia	J	Zelaya	Null
	Jennifer	S	Wallace	Administraçã
	Rames h	K	Naraya n	Null
	Joyce	А	English	Null
	Ahmad	V	Jabbar	Null
	James	Е	Borg	Gerencial

Outer Joins (Junções Externas)

- Um operador similar é a Junção Externa à Direita (*Right Outer Join*), indicada por ► ◄, a qual mantêm as tuplas da segunda relação no resultado.
- Um terceiro operador, Junção Externa Total, (Full Outer Join), indicada por ► ◄, mantém todas as tuplas em ambas as relações, preenchendo quando necessário as tuplas não casadas.
- Esses três operadores fazem parte do padrão SQL-2.

Questões

Estude os exemplos de consulta em Álgebra Relacional do item 7.2 (pág. 60) da apostila de referência.

Responda as questões de 1 à 4 do item 7.3 (pág. 72) existentes na apostila.