

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

EMPREGAD

PNOME MNOME SNOME NSS DATANASC ENDERECO SEXO SALARIO

SUPERVISION
A

TRABALHA-
PARA

NSSSUPER DNUM
 ce c

PROJETO

PNOME PNUMERO PLOCALIZACAO DNUM

ce

CONTROL

A

DEPARTAMENT

DNOME DNUMERO GERNSS GERDATINIC

ce

*

GERENCI

A

DEPENDENT

E ENSS NOMEDEPENDENTE SEXO DATANASC RELACAO

ce

C

TRABALHA- EM

ENSS PNO HORAS

ce

ce

LOCAIS DEPTO

DNUMERO DLOCALIZACAO

ce



Operador SELECT - σ

- Utilizada para selecionar, segundo alguma condição, tuplas de uma relação.

- Exemplos:

- Selecionar os empregados que trabalham para o departamento 4:

$\sigma_{\text{dnum} = 4}(\text{EMPREGADO})$

- Selecionar empregados que tenham salário maior que 3000

$\sigma_{\text{SALÁRIO} > 3000}(\text{EMPREGADO})$

Operador SELECT - σ

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

Operador SELECT - σ

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.

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

Operador SELECT - σ

□ Consulta:

$\sigma_{(NDEP = 4 \text{ AND SALÁRIO} > 2500) \text{ OR } (NDEP = 5 \text{ AND SALÁRIO} > 3000)}$ (EMPREGADO)

□ Resultado:

PNOME	MNOME	SNOME	NSS	DATANASC	ENDEREÇO	SEXO	SALARIO	NSSSUPER	NDEP
Franklin	T	Wong	33344555	08-DEZ-45	R. B, 2	M	4000	88866555 5	5
Jennifer	S	Wallace	98765432	20-JUN-31	Trav. D, 4	F	4300	88866555 5	4
Ramesh	K	Narayan	66688444	15-SET-52	R. E, 5	M	3800	33344555 5	5

Operador SELECT - σ

❑ 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

$$\sigma_{\langle \text{cond}_1 \rangle}(\sigma_{\langle \text{cond}_2 \rangle}(R)) = \sigma_{\langle \text{cond}_1 \rangle \text{ AND } \langle \text{cond}_2 \rangle}(R)$$

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 SNAME, PNAME e SALÁRIO da relação EMPREGADO.

PNAME	MNAME	SNAME	NSS	DATANASC	ENDEREÇO	SEXO	SALARIO	NSSSUPER	NDEP
John	B	Smith	1234567	09-JAN-55	R. A, 1	M	3000	33344555	5
Franklin	T	Wong	3334455	08-DEZ-45	R. B, 2	M	4000	88866555	5
Alícia	J	Zelaya	9998877	19-JUL-58	Av. C, 3	F	2500	98765432	4
Jennifer	S	Wallace	9876543	20-JUN-31	Trav. D, 4	F	4300	88866555	4
Ramesh	K	Naraya	6668844	15-SET-52	R. E, 5	M	3800	33344555	5
Joyce	A	English	4534534	31-JUL-62	R. F, 6	F	2500	33344555	5
Ahmad	V	Jabbar	9879879	29-MAR-59	Av G, 7	M	2500	98765432	4
James	E	Borg	8886655	10-NOV-27	Av H, 8	M	5500	null	3

Operador PROJECT - π

□ Consulta:

π PNAME, SNAME, SALÁRIO (EMPREGADO)

□ Resultado:

PNAME	SNAME	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

π SEXO, SALÁRIO (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
 - ❑ $\langle \text{lista de atributos 1} \rangle \subset \langle \text{lista de atributos 2} \rangle$ então,

$$\pi_{\langle \text{lista1} \rangle} (\pi_{\langle \text{lista2} \rangle} (R)) = \pi_{\langle \text{lista1} \rangle} (R)$$

Seqüência 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	NSSUPER	NDEP
John	B	Smith	12345678	09-JAN-55	R. A, 1	M	3000	333445555	5
Franklin	T	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	A	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

Seqüência 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:

$$\text{DEP5_EMPS} \leftarrow \sigma_{\text{NDEP}=5}(\text{EMPREGADO})$$
$$\text{RESULT} \leftarrow \pi_{\text{PNOME, SNOME, SALÁRIO}}(\text{DEP5_EMPS})$$

Renomeando Atributos

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

$\text{DEP5_EMPS} \leftarrow \sigma_{\text{NDEP}=5}(\text{EMPREGADO})$

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

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

Operações da Teoria dos Conjuntos

- 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  $\leftarrow$   $\sigma_{NDEP=5}$  (EMPREGADO)  
RESULT1  $\leftarrow$   $\pi_{NSS}$  (DEP5_EMPS)  
RESULT2(NSS)  $\leftarrow$   $\pi_{NSSSUPER}$  (DEP5_EMPS)  
RESULT  $\leftarrow$  RESULT1  $\cup$  RESULT2
```

Operações da Teoria dos Conjuntos

□ Os operadores são:

- $R \cup S$ - União (todas as tuplas de R e todas de S)
- $R \cap S$ - Intersecção (todas as tuplas comuns a R e S)
- $R - S$ - Diferença (todas as tuplas de R que não estão em S)
- $R \times S$ – Produto Cartesiano (combinação das tuplas de R com as de S)

Operações da Teoria dos Conjuntos

□ Os operadores são:

- $R \cup S$ - União (todas as tuplas de R e todas de S)
- $R \cap S$ - Intersecção (todas as tuplas comuns a R e S)
- $R - S$ - Diferença (todas as tuplas que não estão em S)
- $R \times S$ - Produto Cartesiano (todas as combinações das tuplas de R com as de S)

**UNIÃO
INTERSECÇÃO
DIFERENÇA**

**PRECISAM QUE R E S SEJAM
COMPATÍVEIS NA UNIÃO**

Operações da Teoria dos Conjuntos

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

Operações da Teoria dos Conjuntos

- Para facilitar, considere o seguinte exemplo:

SNAME	NSS
Smith	12345678
Wong	33344555

X

NSSEMP	NOMEDEPENDENT
33344555	Alice
98765432	Abner
12345678	Alice
12345678	Elizabeth



SNAME	NSS	NSSEMP	NOMEDEPENDENTE
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

Operações da Teoria dos Conjuntos

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

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

$EMP_NOMES \leftarrow \pi_{PNAME, SNAME, NSS}(EMP_FEM)$

$EMP_DEP \leftarrow EMP_NOMES \times DEPENDENTE$

$DEP_ATUAL \leftarrow \sigma_{NSS=ENSS}(EMP_DEP)$

$RESULT \leftarrow \pi_{PNAME, SNAME, NOMEDEPENDENTE}(DEP_ATUAL)$

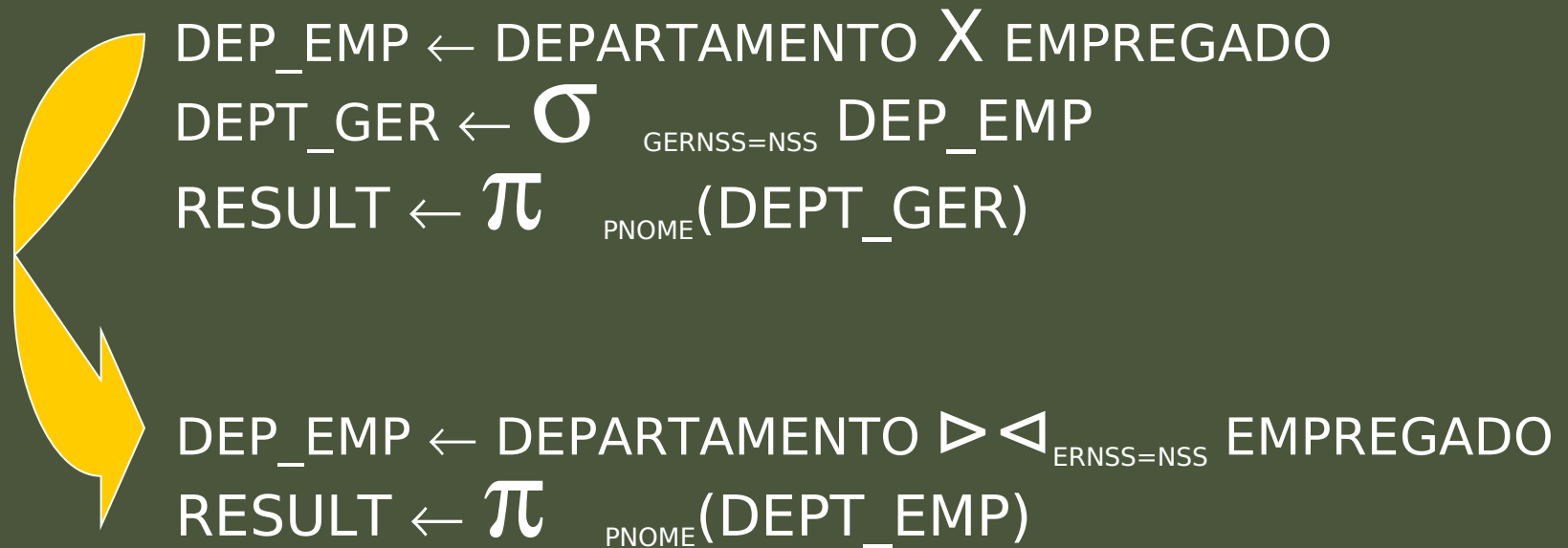
Ver resultado na página 113 do livro do Elmasri & Navathe

Operador Join (\bowtie)

- ❑ O **join** (\bowtie) 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



Operador Join

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

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

$EMP_NOMES \leftarrow \pi_{PNAME, SNAME, NSS} (EMP_FEM)$

$EMP_DEP \leftarrow EMP_NOMES \times DEPENDENTE$

$DEP_ATUAL \leftarrow \sigma_{NSS=ENSS} (EMP_DEP)$

$RESULT \leftarrow \pi_{PNAME, SNAME, NOMEDEPENDENTE} (DEP_ATUAL)$

Operador Join

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

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

$EMP_NOMES \leftarrow \pi_{PNAME, SNAME, NSS} (EMP_FEM)$

$DEP_ATUAL \leftarrow EMP_NOMES \bowtie_{NSS=ENSS} DEPENDENTE$

$RESULT \leftarrow \pi_{PNAME, SNAME, NOMEDEPENDENTE} (DEP_ATUAL)$

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 \leftarrow DEPARTAMENTO $\triangleright \triangleleft$ _{GERNSS=NSS} EMPREGADO

DNOME	DNUMERO	NSSGER	...	PNOME	MNOME	SNOME	NSS	...
Pesquisa	5	333445555	...	Franklin	T	Wong	333445555	...
Administrati	4	987654321	...	Jennifer	S	Wallace	987654321	...
Gerencial	1	888665555	...	James	E	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:
 - $Q \leftarrow R \underset{(lista1), (lista2)}{*} S$, 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 \leftarrow PROJETO *_{DNUM = DNÚMERO} DEPARTAMENTO
 - ou
 - PROJ_DEPT \leftarrow 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 \leftarrow DEPARTAMENTO * LOCAIS_DEPTO

Operador DIVISION \div

- 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:
 - $SMITH \leftarrow \sigma_{PNAME='John' \text{ AND } SNAME='Smith'}(EMPREGADO)$
 - $SMITH_PNO \leftarrow \pi_{PNO}(TRABALHA_EM \div_{NSEMP = NSS} SMITH)$

SMITH_PNO	PNO
	1
	2

Operador DIVISION ÷

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

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
888665555		20

Operador DIVISION ÷

- Finalmente, aplicar a divisão:
 - $NSS_DESEJADOS(NSS) \leftarrow NSS_PNRO \div SMITH_PNO$

NSS_PNRO	NSSEMP	PNO	SMITH_PNO	PNO	NSS_DESEJADO	NSS
123456789		1		1		123456789
123456789		2		2		453453453
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				
888665555		20				

Operador DIVISION \div

- A divisão pode ser escrita em termos dos operadores básicos: π , \times e $-$.
- Sejam duas instâncias de relação $A(x, y)$ e $B(y)$.
- O resultado de $A \div B$ 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 $\langle x, y \rangle$ que não estão em A .
- A seguinte expressão permite fazer isso:
 - $X_{\text{DESQUALIFICADO}} \leftarrow \pi_x ((\pi_x (A) \times B) - A)$

Operador DIVISION \div

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

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 - \mathfrak{S}

Funções de Agregação

- Por exemplo: Recuperar para cada departamento, o número de empregado e sua média salarial.

■ $R(DNO, NRO_EMPS, MÉDIA) \leftarrow DNUM \text{ } \S \text{ COUNT NSS, AVERAGE SALÁRIO (EMPREGADO)}$



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

$R \leftarrow \mathfrak{S} \text{ COUNT_NSS, AVERAGE_SALÁRIO (EMPREGADO)}$

R	COUNT_NSS	AVERAGE_SALÁRIO
	8	3512.5

Clausura Recursiva

- 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.

Clausura Recursiva

- 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

Clausura Recursiva

□ Por exemplo:

- Obter o NSS de todos os empregados diretamente supervisionados por James Borg

- $$\text{BORG_NSS} \leftarrow \pi_{\text{NSS}} (\sigma_{\text{PNOME} = \text{'James'} \text{ AND } \text{SNOME} = \text{'Borg'}} (\text{EMPREGADO}))$$

- $$\text{SUPERVISÃO}(\text{NSS1}, \text{NSS2}) \leftarrow \pi_{\text{NSS}, \text{NSSSUPER}} (\text{EMPREGADO})$$

- $$\text{RESULT1} \leftarrow \pi_{\text{NSS1}} (\text{SUPERVISÃO} \triangleright \triangleleft_{\text{NSS2} = \text{NSS}} \text{BORG_NSS})$$

Clausura Recursiva

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

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 as 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 \bowtie :

- TEMP \leftarrow (EMPREGADO $\bowtie_{NSS=GERNSS}$ DEPARTAMENTO)

- RESULTADO $\leftarrow \pi_{PNAME, MNAME, SNAME, DNAME}$ (TEMP)

RESULTADO	PNAME	MNAME	SNAME	DNAME
	John	B	Smith	Null
	Franklin	T	Wong	Pesquisa
	Alícia	J	Zelaya	Null
	Jennifer	S	Wallace	Administração
	Ramesh	K	Narayan	Null
	Joyce	A	English	Null
	Ahmad	V	Jabbar	Null
	James	E	Borg	Gerencial

Outer Joins (Junções Externas)

- ❑ Um operador similar é a Junção Externa à Direita (*Right Outer Join*), indicada por $\triangleright \triangleleft =$, a qual mantêm as tuplas da segunda relação no resultado.
- ❑ Um terceiro operador, Junção Externa Total, (*Full Outer Join*), indicada por $= \triangleright \triangleleft =$, 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.