

Operações relacionais e Álgebra relacional

Introdução

A discussão sobre algumas operações básicas de álgebra relacional realizada a seguir considera um banco de dados composto pelas seguintes relações:

funcionário (NrMatric, NmFunc, DtAdm, Sexo, CdCargo, CdDepto)

cargo (CdCargo, NmCargo, VrSalario)

depto (CdDepto, NmDepto, Ramal)

Os atributos sublinhados com linha contínua indicam a chave primária de cada relação enquanto que aqueles sublinhados com linha pontilhada constituem chaves estrangeiras.

cargo

<u>CdCargo</u>	NmCargo	VrSalário
C1	COZINHEIRA	350
C3	AUX. ESCRITÓRIO	450
C7	VIGIA	400
C2	MECANICO	750
C5	GERENTE	2300
C4	ESCRITURARIO	600

depto

<u>CdDepto</u>	NmDepto	Ramal
D1	ADMINISTRACAO	221
D2	OFICINA	235
D3	SERVICOS GERAIS	243
D4	VENDAS	258

funcionário

<u>NrMatric</u>	NmFunc	DtAdm	Sexo	<u>CdCargo</u>	<u>CdDepto</u>
1001	JOAO SAMPAIO	10/08/93	M	C2	D2
1004	LUCIO TORRES	02/03/94	M	C2	D2
1034	ROBERTO PEREIRA	23/05/92	M	C3	D1
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1
1029	RUTH DE SOUZA	05/01/92	F	C1	D3
1095	MARIA DA SILVA	03/09/92	F	C4	D1
1023	LUIZ DE ALMEIDA	12/01/93	M	C2	D2
1042	PEDRO PINHEIRO	29/07/94	M	C4	D1
1048	ANA SILVEIRA	01/06/93	F	C5	D1
1015	PAULO RODRIGUES	17/08/92	M	C2	D2

Estamos interessados em obter informações armazenadas nesse banco de dados, e para isso deveremos formular expressões em álgebra relacional combinando apenas algumas operações primitivas clássicas, que serão apresentadas uma a uma de forma simplificada em termos de seu significado prático e notação. Toda operação relacional opera (age) sobre um ou mais conjuntos de dados e fornece como resultado um novo conjunto. Devido a essa característica, podemos combinar mais de uma operação relacional em uma única expressão algébrica, fazendo com que o resultado de uma operação seja utilizado como entrada para outra operação, aumentando com isso grandemente o poder dessa linguagem de consulta. Suponha que, inicialmente, precisamos obter o nome completo de todos os funcionários de nosso banco de dados. Para isso será necessário executar uma operação chamada Projeção.

Projeção

Geralmente indicada na literatura por \mathbf{p} (a letra grega ρ) produz um conjunto onde há um elemento para cada elemento do conjunto de entrada, sendo que a estrutura dos membros do conjunto resultante é definida nos argumentos da operação. Pode ser entendida como uma operação que filtra as colunas de uma tabela. Por operar sobre apenas um conjunto de entrada, a projeção é classificada como uma *operação unária*.

Ex.: \mathbf{p}_{NmFunc} (funcionário)

essa expressão produz um conjunto contendo um elemento para cada funcionário, e cada elemento contém apenas a informação referente a NmFunc da relação funcionário original.

funcionário

NrMatric	NmFunc	DtAdm	Sexo	CdCargo	CdDepto
1001	JOAO SAMPAIO	10/08/93	M	C2	D2
1004	LUCIO TORRES	02/03/94	M	C2	D2
1034	ROBERTO PEREIRA	23/05/92	M	C3	D1
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1
1029	RUTH DE SOUZA	05/01/92	F	C1	D3
1095	MARIA DA SILVA	03/09/92	F	C4	D1
1023	LUIZ DE ALMEIDA	12/01/93	M	C2	D2
1042	PEDRO PINHEIRO	29/07/94	M	C4	D1
1048	ANA SILVEIRA	01/06/93	F	C5	D1
1015	PAULO RODRIGUES	17/08/92	M	C2	D2



funcionário'

NmFunc
JOAO SAMPAIO
LUCIO TORRES
ROBERTO PEREIRA
JOSE NOGUEIRA
RUTH DE SOUZA
MARIA DA SILVA
LUIZ DE ALMEIDA
PEDRO PINHEIRO
ANA SILVEIRA
PAULO RODRIGUES

Agora estamos interessados em identificar todos os funcionários de sexo masculino existentes no banco de dados. É uma situação que não podemos resolver com projeções apenas, uma vez que deveremos descartar elementos do conjunto inicial. Para casos desse tipo existe uma operação relacional chamada Seleção.

Seleção (ou Restrição)

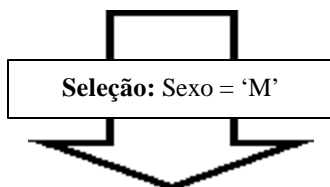
Indicada por σ (a letra grega *sigma*), é uma operação que para um conjunto inicial fornecido como argumento, produz um subconjunto estruturalmente idêntico, mas apenas com os elementos do conjunto original que atendem a uma determinada condição (também chamada de *predicado*). A seleção pode ser entendida como uma operação que filtra as linhas de uma tabela, e é também uma operação unária, já que opera sobre um único conjunto de dados.

Ex.: $\sigma_{\text{Sexo} = 'M'}$ (funcionário)

produz o conjunto dos elementos de funcionário que atendem ao predicado [Sexo = 'M'], ou seja, representa um subconjunto dos funcionários para o qual essa condição é avaliada como verdadeira.

funcionário

NrMatric	NmFunc	DtAdm	Sexo	CdCargo	CdDepto
1001	JOAO SAMPAIO	10/08/93	M	C2	D2
1004	LUCIO TORRES	02/03/94	M	C2	D2
1034	ROBERTO PEREIRA	23/05/92	M	C3	D1
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1
1029	RUTH DE SOUZA	05/01/92	F	C1	D3
1095	MARIA DA SILVA	03/09/92	F	C4	D1
1023	LUIZ DE ALMEIDA	12/01/93	M	C2	D2
1042	PEDRO PINHEIRO	29/07/94	M	C4	D1
1048	ANA SILVEIRA	01/06/93	F	C5	D1
1015	PAULO RODRIGUES	17/08/92	M	C2	D2

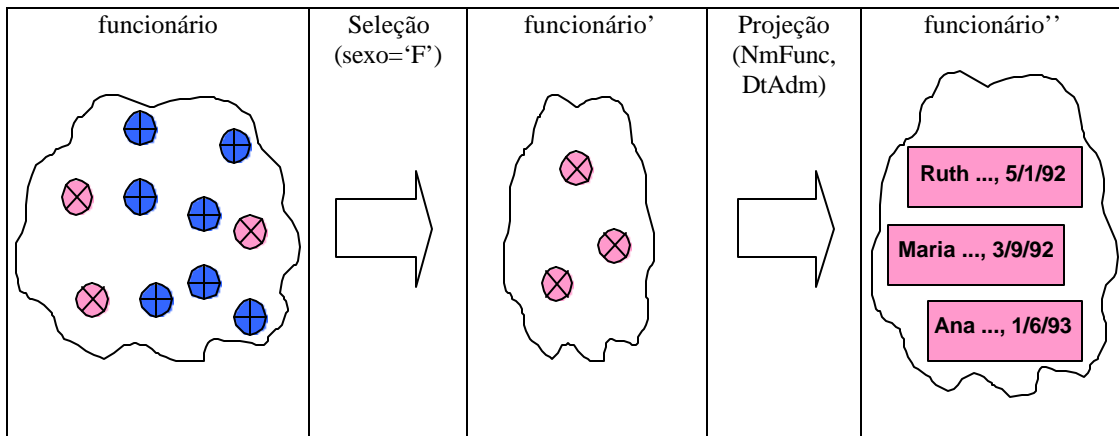


funcionário'

NrMatric	NmFunc	DtAdm	Sexo	CdCargo	CdDepto
1001	JOAO SAMPAIO	10/08/93	M	C2	D2
1004	LUCIO TORRES	02/03/94	M	C2	D2
1034	ROBERTO PEREIRA	23/05/92	M	C3	D1
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1
1023	LUIZ DE ALMEIDA	12/01/93	M	C2	D2
1042	PEDRO PINHEIRO	29/07/94	M	C4	D1
1015	PAULO RODRIGUES	17/08/92	M	C2	D2

No caso de quisermos descobrir o nome completo e a data de admissão de todos os funcionários do sexo feminino existentes na empresa, será necessário combinar uma projeção com uma seleção. Isso porque se decidirmos projetar as colunas desejadas diretamente a partir da relação funcionário, estaremos considerando também os elementos do sexo masculino, o que não queremos. Como a projeção não permite descartar linhas, apenas colunas, deveremos fornecer a essa operação o subconjunto resultante de uma filtragem (seleção) da relação de funcionários original, como mostram as duas figuras a seguir, que representam as relações e as operações de duas maneiras diferentes.

Operações relacionais e Álgebra relacional



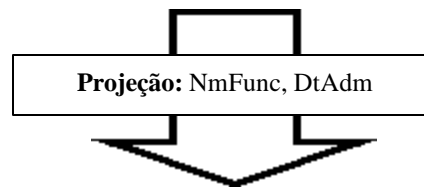
funcionário

NrMatric	NmFunc	DtAdm	Sexo	CdCargo	CdDepto
1001	JOAO SAMPAIO	10/08/93	M	C2	D2
1004	LUCIO TORRES	02/03/94	M	C2	D2
1034	ROBERTO PEREIRA	23/05/92	M	C3	D1
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1
1029	RUTH DE SOUZA	05/01/92	F	C1	D3
1095	MARIA DA SILVA	03/09/92	F	C4	D1
1023	LUIZ DE ALMEIDA	12/01/93	M	C2	D2
1042	PEDRO PINHEIRO	29/07/94	M	C4	D1
1048	ANA SILVEIRA	01/06/93	F	C5	D1
1015	PAULO RODRIGUES	17/08/92	M	C2	D2



funcionário'

NrMatric	NmFunc	DtAdm	Sexo	CdCargo	CdDepto
1029	RUTH DE SOUZA	05/01/92	F	C1	D3
1095	MARIA DA SILVA	03/09/92	F	C4	D1
1048	ANA SILVEIRA	01/06/93	F	C5	D1



funcionário''

NmFunc	DtAdm
RUTH DE SOUZA	05/01/92
MARIA DA SILVA	03/09/92
ANA SILVEIRA	01/06/93

Assim, a expressão que atende nossos objetivos nesse caso é

P NmFunc, DtAdm (**S** Sexo = 'F' (funcionário))

cabendo observar que devido ao aninhamento das operações está implícito que primeiro será executada a seleção e depois a projeção, sendo que nesse exemplo não poderíamos inverter essa ordem (você poderia explicar porquê?).

Por esse motivo, dizemos que a álgebra relacional é uma *linguagem procedural*, já que requer alguma definição quanto à ordem em que as operações serão realizadas. Linguagens em que apenas mencionamos o resultado desejado, sem fazer menção alguma à forma como isso deve ser feito são chamadas de *linguagens não-procedurais*.

Suponha agora que precisamos obter o nome completo, a data de admissão e o salário de cada funcionário cadastrado. Para essa consulta temos um fato novo, que é a referência a colunas de mais de uma tabela, uma vez que o nome e a data de admissão fazem parte da relação funcionário, enquanto que o salário existe apenas em cargos. Isso é problemático, pois as duas operações que conhecemos até o momento são unárias, e temos necessidade de combinar os dados de mais de uma relação. Para situações como essa existe uma operação chamada Produto Cartesiano.

Produto cartesiano

A notação geralmente adotada (na forma 'conjunto1 x conjunto2') para representar essa *operação binária* indica bem a sua natureza: o resultado do produto cartesiano de duas tabelas é uma terceira tabela contendo todas as combinações possíveis entre os elementos das tabelas originais. Essa tabela resultante possuirá um número de colunas que é igual à soma das quantidades de colunas das duas tabelas iniciais, e um número de linhas igual ao produto do número de suas linhas. Portanto, se fizermos o produto cartesiano de uma tabela A que possua 4 colunas e 10 linhas com uma tabela B onde existem 3 colunas e 7 linhas, a tabela resultante terá $4+3=7$ colunas e $10*7=70$ linhas. Assim, cada linha dessa tabela corresponderá à concatenação de uma linha da primeira tabela com uma linha da segunda. O produto cartesiano não é muito usado como um fim em si mesmo, ou seja, dificilmente estaremos interessados em saber quais são todas as combinações possíveis entre as linhas de duas tabelas, pois a utilidade prática desse tipo de conhecimento é muito discutível. Entretanto, é a única forma primitiva de que dispomos para fundir informações de duas tabelas heterogêneas para posterior processamento. Nesse caso, tipicamente será necessário executar uma Seleção sobre o resultado do Produto Cartesiano, de maneira a descartar as combinações inválidas entre as linhas das tabelas originais.

Ex.: **P** NmFunc, DtAdm, VrSalário (**S** funcionário.CdCargo = cargo.CdCargo (funcionário x cargo))

observe que primeiro é produzido o produto cartesiano correspondente a todas as combinações possíveis entre funcionários e cargos. Essa relação vai conter linhas onde um funcionário estará associado a cargos que não são o seu, e devemos então aplicar um filtro (uma seleção) para gerar um subconjunto apenas com as combinações logicamente válidas (aquelas em que a chave estrangeira CdCargo de funcionário tem valor igual à chave primária CdCargo de cargo). Como temos nesse subconjunto duas colunas com o mesmo nome (CdCargo que veio de funcionário e CdCargo proveniente de cargo), sempre que precisarmos mencionar uma delas será necessário especificar exatamente a qual das duas

Operações relacionais e Álgebra relacional

colunas estamos nos referindo, senão teremos uma situação ambígua, formalmente inaceitável. Dizemos, nesse caso, que é necessário qualificar a coluna, e isso é feito escrevendo o nome da relação original antes do nome da coluna, separando-os por um ponto, ou seja, <nome-da-relação>.<nome-da-coluna>. É por esse motivo que escrevemos o predicado da seleção como sendo *funcionário.CdCargo = cargo.CdCargo*. Finalmente, a projeção é realizada a partir desse subconjunto, fornecendo os dados inicialmente desejados.

funcionário					
NrMatric	NmFunc	DtAdm	Sexo	CdCargo	CdDepto
1001	JOAO SAMPAIO	10/08/93	M	C2	D2
1004	LUCIO TORRES	02/03/94	M	C2	D2
1034	ROBERTO PEREIRA	23/05/92	M	C3	D1
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1
1029	RUTH DE SOUZA	05/01/92	F	C1	D3
1095	MARIA DA SILVA	03/09/92	F	C4	D1
1023	LUIZ DE ALMEIDA	12/01/93	M	C2	D2
1042	PEDRO PINHEIRO	29/07/94	M	C4	D1
1048	ANA SILVEIRA	01/06/93	F	C5	D1
1015	PAULO RODRIGUES	17/08/92	M	C2	D2

cargo		
CdCargo	NmCargo	VrSalário
C1	COZINHEIRA	350
C3	AUX. ESCRIT.	450
C7	VIGIA	400
C2	MECANICO	750
C5	GERENTE	2300
C4	ESCRITURARIO	600



(continua na próxima página)

Operações relacionais e Álgebra relacional

funcionário_cargo

NrMatric	NmFunc	DtAdm	Sexo	CdCargo	CdDepto	CdCargo	NmCargo	VrSalario
1001	JOAO SAMPAIO	10/08/93	M	C2	D2	C1	COZINHEIRA	350
1001	JOAO SAMPAIO	10/08/93	M	C2	D2	C3	AUX. ESCRITÓRIO	450
1001	JOAO SAMPAIO	10/08/93	M	C2	D2	C7	VIGIA	400
1001	JOAO SAMPAIO	10/08/93	M	C2	D2	C2	MECANICO	750
1001	JOAO SAMPAIO	10/08/93	M	C2	D2	C5	GERENTE	2300
1001	JOAO SAMPAIO	10/08/93	M	C2	D2	C4	ESCRITURARIO	600
...
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1	C1	COZINHEIRA	350
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1	C3	AUX. ESCRITÓRIO	450
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1	C7	VIGIA	400
...
1015	PAULO RODRIGUES	17/08/92	M	C2	D2	C4	ESCRITURARIO	600

Seleção: funcionário.CdCargo = cargo.CdCargo

funcionário_cargo'

NrMatric	NmFunc	DtAdm	Sexo	CdCargo	CdDepto	CdCargo	NmCargo	VrSalario
1001	JOAO SAMPAIO	10/08/93	M	C2	D2	C2	MECANICO	750
1004	LUCIO TORRES	02/03/94	M	C2	D2	C2	MECANICO	750
1034	ROBERTO PEREIRA	23/05/92	M	C3	D1	C3	AUX. ESCRITÓRIO	450
1021	JOSE NOGUEIRA	10/11/94	M	C3	D1	C3	AUX. ESCRITÓRIO	450
1029	RUTH DE SOUZA	05/01/92	F	C1	D3	C1	COZINHEIRA	350
1095	MARIA DA SILVA	03/09/92	F	C4	D1	C4	ESCRITURARIO	600
1023	LUIZ DE ALMEIDA	12/01/93	M	C2	D2	C2	MECANICO	750
1042	PEDRO PINHEIRO	29/07/94	M	C4	D1	C4	ESCRITURARIO	600
1048	ANA SILVEIRA	01/06/93	F	C5	D1	C5	GERENTE	2300
1015	PAULO RODRIGUES	17/08/92	M	C2	D2	C2	MECANICO	750

Projeção: NmFunc, DtAdm, VrSalário

(continua na próxima página)

funcionário_cargo''

NmFunc	DtAdm	VrSalario
JOAO SAMPAIO	10/08/93	750
LUCIO TORRES	02/03/94	750
ROBERTO PEREIRA	23/05/92	450
JOSE NOGUEIRA	10/11/94	450
RUTH DE SOUZA	05/01/92	350
MARIA DA SILVA	03/09/92	600
LUIZ DE ALMEIDA	12/01/93	750
PEDRO PINHEIRO	29/07/94	600
ANA SILVEIRA	01/06/93	2300
PAULO RODRIGUES	17/08/92	750

Caso desejemos obter uma variação ligeiramente diferente dessa consulta, acrescentando a restrição de que precisamos dos dados apenas dos funcionários do sexo masculino, teríamos a seguinte expressão algébrica:

$\rho_{NmFunc, DtAdm, VrSalário}(\sigma_{funcionário.CdCargo = cargo.CdCargo \wedge Sexo = 'M'}(funcionário \times cargo))$
onde o símbolo \wedge presente no predicado representa o conectivo lógico “E”.

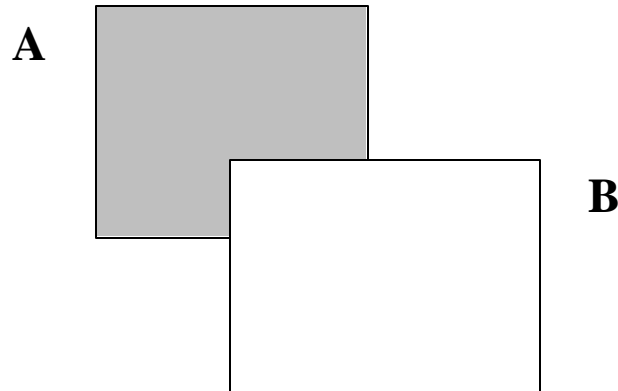
Portanto, com apenas 3 operações relacionais básicas foi possível extrair do banco de dados de exemplo várias informações importantes, representativas de uma grande parcela das consultas que um sistema gerenciador de bancos de dados deve processar.

As consultas realizadas foram:

1. obter o nome completo de todos os funcionários;
2. identificar todos os funcionários do sexo masculino;
3. obter o nome completo e a data de admissão de todos os funcionários do sexo feminino;
4. obter o nome completo, a data de admissão e o salário de todos os funcionários;
5. descobrir o nome completo, a data de admissão e o salário de todos os funcionários do sexo masculino.

Diferença entre conjuntos: $A - B$

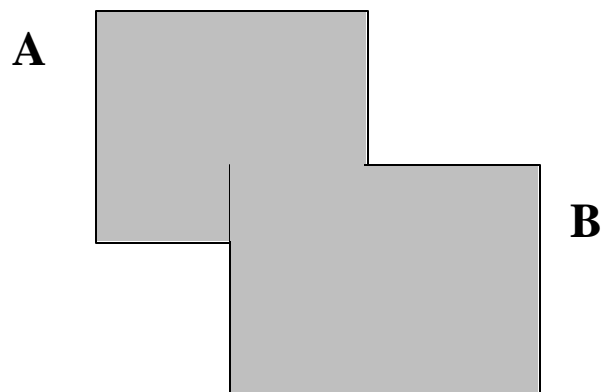
É uma operação primitiva que requer como operandos duas tabelas união-compatíveis, ou seja, estruturalmente idênticas. O resultado é uma tabela que possui todas as linhas que existem na primeira tabela e não existem na segunda.



Observe que $A - B$ é diferente de $B - A$.

União: $A \cup B$

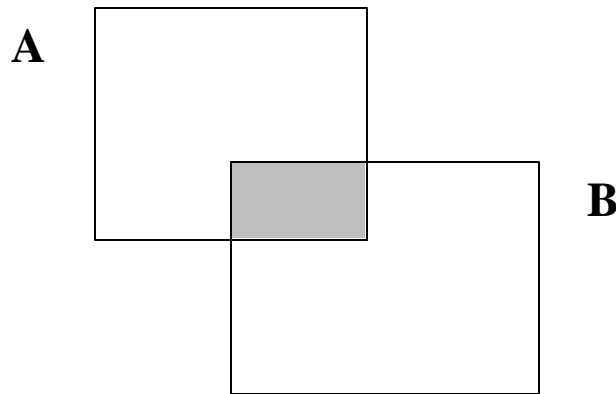
Esta operação primitiva também requer como operandos tabelas união-compatíveis. Produz como resultado uma tabela que contém todas as linhas da primeira tabela seguidas de todas as linhas da segunda tabela. A tabela resultante possui a mesma quantidade de colunas que as tabelas originais, e tem um número de linhas que é no máximo igual à soma das linhas das tabelas fornecidas como operandos, já que as linhas que são comuns a ambas as tabelas aparecem uma única vez no resultado.



Observe que $A \cup B = B \cup A$.

Intersecção: $A \cap B$

Esta é uma operação adicional que produz como resultado uma tabela que contém, sem repetições, todos os elementos que são comuns às duas tabelas fornecidas como operandos. As tabelas devem ser união-compatíveis.



O mesmo efeito pode ser obtido fazendo-se uma combinação de diferenças entre conjuntos

$$A \cap B = A - (A - B)$$

ou com uniões e diferenças

$$A \cap B = (A \cup B) - (A - B) - (B - A) \quad \text{ou ainda}$$

$$A \cap B = (A \cup B) - ((A - B) \cup (B - A))$$

Junção: $A \bowtie B$

É uma operação que produz uma combinação entre as linhas de uma tabela com as linhas correspondentes de outra tabela, sendo em princípio correspondente a uma seleção pelos atributos de relacionamento sobre um produto cartesiano dessas tabelas:

$$A \bowtie B = \sigma_{A.chave1 = B.chave2} (A \times B)$$

A operação de junção foi criada justamente porque esse tipo de combinação de tabelas é de uso muito comum, facilitando com isso a escrita de expressões. A tabela resultante de uma junção tem todas as colunas da primeira tabela e todas da segunda tabela. Isso faz com que os valores dos campos utilizados como critério para a correspondência entre as linhas apareça duplicado, já que um vem da primeira tabela e outro da segunda. Existe uma variação da junção, chamada *junção natural*, que fornece o mesmo resultado, mas sem essa repetição de valores: uma das colunas correspondentes aos atributos de relacionamento é descartada.

Renomeação: $\rho_{\langle \text{nov}_o_nome \rangle} (A)$

Geralmente representada pelo símbolo ρ (a letra grega *rho* minúscula), esta operação unária primitiva redefine o nome de uma tabela em um determinado contexto. É útil para auto-

relacionamentos, onde precisamos fazer a junção de uma tabela com ela mesma, e nesse caso cada versão da tabela precisa receber um nome diferente da outra.

Divisão: $A \div B$

É uma operação adicional que produz como resultado a projeção de todos os elementos da primeira tabela que se relacionam com todos os elementos da segunda tabela.

Ex.: queremos saber os nomes dos departamentos que possuem todos os cargos:

$$\pi_{NmDepto, CdCargo} (depto \mid x \mid funcionario) \div \pi_{CdCargo} (cargo)$$

Atribuição: $variável \leftarrow A$

Permite que o conteúdo de uma tabela seja atribuído (colocado) em uma variável especial, oferecendo a possibilidade de um tratamento até certo ponto algorítmico para algumas seqüências de operações.

Resumo

<i>Símbolo</i>	<i>Operação</i>	<i>Sintaxe</i>	<i>Tipo</i>
s	Seleção / Restrição	$\sigma_{condição} (Relação)$	Primitiva
p	Projeção	$\pi_{expressões} (Relação)$	Primitiva
È	União	Relação1 \cup Relação2	Primitiva
Ç	Intersecção	Relação1 \cap Relação2	Adicional
-	Diferença de conjuntos	Relação1 - Relação2	Primitiva
x	Produto cartesiano	Relação1 x Relação2	Primitiva
 x 	Junção	Relação1 x Relação2	Adicional
,	Divisão	Relação1 \div Relação2	Adicional
r	Renomeação	$\rho_{nome} (Relação)$	Primitiva
↵	Atribuição	variável \leftarrow Relação	Adicional

Exercícios

Com base nas tabelas Funcionário, Cargo e Depto apresentadas neste material, elaborar as expressões da álgebra relacional que obtenham:

- 1) Todos os funcionários do departamento 'D1'.
- 2) O nome e a matrícula de todos os funcionários do departamento 'D1'.
- 3) A matrícula e o nome do respectivo departamento de todos os funcionários.
- 4) O nome dos funcionários que ganham mais de \$500.
- 5) O ramal do funcionário 'ANA SILVEIRA'.
- 6) Os nomes de todos os funcionários com cargo de 'MECANICO'.
- 7) Os nomes de todos os funcionários que trabalham no mesmo departamento que 'JOSE NOGUEIRA'.
- 8) Os nomes dos departamentos que possuem tanto funcionários como funcionárias.