

Universidade Presbiteriana Mackenzie

Banco de Dados - Aula 08

Álgebra Relacional

Profa. Elisângela Botelho Gracias

Faculdade de Computação e Informática



Introdução

- Uma linguagem de consulta é uma linguagem na qual um usuário requisita informações do banco de dados
- Essas linguagens são tipicamente de mais alto nível do que as linguagens de programação comuns

Introdução

- A álgebra relacional é uma linguagem de consulta para bancos de dados, onde as variáveis são relações e os operadores são operadores relacionais

Introdução

- Conceitos na álgebra relacional:
 - Relação: é uma tabela
 - Tuplas: são as linhas da tabela
 - Atributos: são as colunas da tabela



Introdução

- A álgebra relacional é fechada, ou seja, o resultado de uma ou mais operações relacionais é uma nova relação
- O resultado de uma operação pode ser utilizado como dado de entrada de outra

Introdução

- As operações da álgebra relacional podem ser divididas em 2 grupos:
 - Um grupo consiste de operações desenvolvidas especificamente para banco de dados relacionais, que inclui **SELEÇÃO**, **PROJEÇÃO** e **JUNÇÃO**, dentre outras
 - O outro grupo são as operações da teoria de conjuntos da matemática: **UNIÃO**, **INTERSEÇÃO**, **DIFERENÇA** e **PRODUTO CARTESIANO**

Introdução

- Considere a seguinte relação para os exemplos das operações SELECT e PROJECT, a seguir:

EMPREGADO = {NOME, SOBREN, RG, DATAN,
ENDEREÇO, SEX, RG_SUPER, NROD}

- RG é a chave primária desta relação
- RG_SUPER é chave estrangeira que referencia o atributo RG da própria relação Empregado

NOME	SOBREN	<u>RG</u>	DATAN	ENDEREÇO	SEX	SALARIO	RG SUPER	NROD
João	Silva	123456789	09.01.55	Rua 15 Novembro, 731, São Carlos	M	3000	333445555	5
Francisco	Souza	333445555	08.12.45	Rua Jorge Assef, 100, S.Paulo	M	4000	888665555	5
Alice	Fernandes	999887777	18.03.75	Rua9 de Julho, 535, São Carlos	F	2500	987654321	4
Joana	Pereira	987654321	20.06.31	Rua Tiradentes, 291, Rib.Preto	F	4300	888665555	4
Rodolfo	Nogueira	666884444	15.09.52	Av.São Carlos,1005, São Carlos	M	3800	333445555	5

Operações de Banco de Dados Relacional

- Operações de Banco de Dados Relacional
 - Operação SELECT
 - Operação PROJECT
 - Sequências de Operações e Renomeação de Atributos

Operações de Banco de Dados Relacional

- Operações de Banco de Dados Relacional
 - Operação SELECT
 - Operação PROJECT
 - Sequências de Operações e Renomeação de Atributos

Operação SELECT

- A operação SELECT é utilizada para selecionar um subconjunto de tuplas numa relação que satisfaça uma condição de seleção

Operação SELECT

- Utilizando a Linguagem SQL, a seguinte consulta - selecionar o subconjunto de tuplas de EMPREGADO cujo salário é maior do que R\$ 3000,00 - seria:

```
SELECT *  
FROM EMPREGADO  
WHERE (SALARIO > 3000);
```

Operação SELECT

- A mesma consulta dada anteriormente, mas agora utilizando a Álgebra Relacional, seria (o resultado é mostrado, a seguir):

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

- Resultado

NOME	SOBREN	<u>RG</u>	DATAN	ENDEREÇO	SEX	SALARIO	RG SUPER	NROD
Francisco	Souza	333445555	08.12.45	Rua Jorge Assef, 100, S.Paulo	M	4000	888665555	5
Joana	Pereira	987654321	20.06.31	Rua Tiradentes, 291, Rib.Preto	F	4300	888665555	4
Rodolfo	Nogueira	666884444	15.09.52	Av.São Carlos,1005, São Carlos	M	3800	333445555	5

Operação SELECT

- Em geral, a operação SELECT é denotada por:

$\sigma_{\langle \text{condição da seleção} \rangle}$ (Nome da Relação)

- onde o símbolo σ (sigma) é usado para denotar o operador SELECT e a $\langle \text{condição de seleção} \rangle$ é uma expressão booleana especificada nos atributos da relação

Operação SELECT

- A relação resultante da operação SELECT tem os mesmos atributos da relação especificada

Operação SELECT

- Exemplo: selecionar as tuplas para todos os empregados que ou trabalham no departamento 4 e recebem mais de 4000, ou trabalham no departamento 5 e recebem mais de 3000:

$\sigma_{((\text{NROD}=4) \text{ AND } (\text{SALARIO}>4000)) \text{ OR } ((\text{NROD}=5) \text{ AND } (\text{SALARIO}>3000))} (\text{EMPREGADO})$

Operação SELECT

- Utilizando a Linguagem SQL, a consulta anterior seria:

SELECT *

FROM EMPREGADO

WHERE ((NROD = 4) AND (SALARIO > 4000))

OR ((NROD = 5) AND (SALARIO > 3000));

Operação SELECT

- Resultado

NOME	SOBREN	<u>RG</u>	DATAN	ENDEREÇO	SEX	SALARIO	RG SUPER	NROD
Francisco	Souza	333445555	08.12.45	Rua Jorge Assef, 100, S.Paulo	M	4000	888665555	5
Joana	Pereira	987654321	20.06.31	Rua Tiradentes, 291, Rib.Preto	F	4300	888665555	4
Rodolfo	Nogueira	666884444	15.09.52	Av.São Carlos,1005, São Carlos	M	3800	333445555	5

Operação SELECT

- A operação SELECT é comutativa, ou seja:

$$\sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(R)) = \sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond1} \rangle}(R))$$

- Assim, uma sequência de SELECT's podem ser aplicadas em qualquer ordem

Operações de Banco de Dados Relacional

- Operações de Banco de Dados Relacional
 - Operação SELECT
 - Operação PROJECT
 - Sequências de Operações e Renomeação de Atributos

Operação PROJECT

- A operação PROJECT seleciona certas colunas da tabela e descarta as outras colunas
- Exemplo: para listar o nome, o sobrenome e o salário de cada empregado, utilizando a **Álgebra Relacional**, pode-se usar a operação PROJECT da seguinte forma:

$\pi_{\text{NOME, SOBREN, SALARIO}}$ (EMPREGADO)

Operação PROJECT

- Resultado

NOME	SOBREN	SALARIO
João	Silva	3000
Francisco	Souza	4000
Alice	Fernandes	2500
Joana	Pereira	4300
Rodolfo	Nogueira	3800

Operação PROJECT

- Utilizando a **Linguagem SQL**, para a consulta anterior, seria:

```
SELECT NOME, SOBREN, SALARIO  
FROM EMPREGADO;
```

Operação PROJECT

- A forma geral da operação PROJECT é:

$\pi_{\langle \text{lista de atributos} \rangle}$ (Nome da Relação)

Onde:

- π (pi) é o símbolo usado para representar a operação PROJECT
- $\langle \text{lista de atributos} \rangle$ é uma lista de atributos da relação especificada

Operação PROJECT

- A relação resultante tem somente os atributos especificados em <lista de atributos> e na mesma ordem que aparecem na lista

Operação PROJECT

- A operação PROJECT remove implicitamente qualquer tupla duplicada. Assim, o resultado de uma operação PROJECT é um conjunto de tuplas e, portanto, uma relação válida
- O número de tuplas em uma relação resultante de uma operação PROJECT é sempre menor ou igual ao número de tuplas da relação original

Operação PROJECT

- Se a lista de projeção inclui uma chave da relação, a relação resultante tem o mesmo número de tuplas que o original
- A comutatividade não é mantida em PROJECT

Operações de Banco de Dados Relacional

- Operações de Banco de Dados Relacional
 - Operação SELECT
 - Operação PROJECT
 - Sequências de Operações e Renomeação de Atributos

Sequências de Operações e Renomeação de Atributos

- As operações podem ser aplicadas de forma aninhada, usando uma única expressão da álgebra relacional
- Exemplo: recupere o nome, sobrenome e o salário de todos os empregados que trabalham no departamento cujo código é 5:

$\pi_{\text{NOME}, \text{SOBREN}, \text{SALARIO}} (\sigma_{(\text{NROD} = 5)} (\text{EMPREGADO}))$

Sequências de Operações e Renomeação de Atributos

- Resultado

NOME	SOBREN	SALARIO
João	Silva	3000
Francisco	Souza	4000
Rodolfo	Nogueira	3800

Sequências de Operações e Renomeação de Atributos

- Pode-se, também, criar relações resultantes intermediárias, dando um nome para cada relação intermediária:

$\text{DEP5EMPS} \leftarrow \sigma_{(\text{NROD} = 5)} (\text{EMPREGADO})$

$\text{RESULT} \leftarrow \pi_{\text{NOME, SOBREN, SALARIO}} (\text{DEP5EMPS})$

Sequências de Operações e Renomeação de Atributos

- Pode-se, também, utilizar esta técnica para renomear os atributos nas relações intermediárias e resultantes
- Isso pode ser útil na conexão com operações mais complexas tais como UNION e JOIN

$TEMP \leftarrow \sigma_{(NROD = 5)}(EMPREGADO)$

$R(FIRSTNAME, LASTNAME, SALARY) \leftarrow \pi_{NOME, SOBREN, SALARIO}(TEMP)$

Sequências de Operações e Renomeação de Atributos

- 1º Resultado (TEMP)

NOME	SOBREN	<u>RG</u>	DATAN	ENDEREÇO	SEX	SALARIO	RG SUPER	NROD
João	Silva	123456789	09.01.55	Rua 15 Novembro, 731, São Carlos	M	3000	333445555	5
Francisco	Souza	333445555	08.12.45	Rua Jorge Assef, 100, S.Paulo	M	4000	888665555	5
Rodolfo	Nogueira	666884444	15.09.52	Av.São Carlos,1005, São Carlos	M	3800	333445555	5

Sequências de Operações e Renomeação de Atributos

- 2º Resultado (R)

FIRSTNAME	LASTNAME	SALARY
João	Silva	3000
Francisco	Souza	4000
Rodolfo	Nogueira	3800

Operação PROJECT

- Utilizando a **Linguagem SQL**, para a consulta anterior - recupere o nome, sobrenome e o salário de todos os empregados que trabalham no departamento cujo código é 5 - seria:

```
SELECT NOME, SOBREN, SALARIO  
FROM EMPREGADO  
WHERE (NROD = 5);
```

Operações da Teoria de Conjuntos

- Operações da Teoria de Conjuntos
 - Union
 - Intersection
 - Difference
 - Produto Cartesiano
 - Join



Operações da Teoria de Conjuntos

- O próximo grupo de operações da álgebra relacional são operações matemáticas padrões em conjuntos
- Elas se aplicam ao modelo relacional, pois uma relação é definida como um conjunto de tuplas e pode ser usada para processar as tuplas em duas relações como conjuntos

Operações da Teoria de Conjuntos

- Para a utilização dessas operações em banco de dados relacionais deve-se garantir que o resultado da aplicação dessas operações em duas relações gere uma terceira relação também válida
- Para isso, as duas relações envolvidas devem possuir o mesmo tipo de tuplas; esta condição é chamada **compatibilidade de união**

Operações da Teoria de Conjuntos

- Duas relações União-Compatíveis

R

N	S
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

S

Nome	Sobrenome
John	Smith
Ricardo	Browne
Susan	Yao
Francis	Johnson
Ramesh	Shah

Operações da Teoria de Conjuntos

- Union
- Intersection
- Difference
- Produto Cartesiano
- Join



Operações da Teoria de Conjuntos

- Union
- Intersection
- Difference
- Produto Cartesiano
- Join



Union

- Considerando as 2 relações união-compatíveis R e S, vistas anteriormente:
 - UNION, denotada por $R \cup S$, tem como resultado uma relação que inclui todas as tuplas que estão em R ou em S ou em ambas R e S. Tuplas duplicadas são eliminadas

Union

- Resultado da interseção de $R \cup S$

N	S
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert
John	Smith
Ricardo	Browne
Francis	Johnson

Operações da Teoria de Conjuntos

- Union
- Intersection
- Difference
- Produto Cartesiano
- Join



Intersection

- Considerando as 2 relações união-compátíveis R e S , vistas anteriormente:
 - INTERSECTION, denotada por $R \cap S$, tem como resultado uma relação que inclui todas as tuplas que estão em ambas R e S

Intersection

- Resultado da união de $R \cap S$

N	S
Susan	Yao
Ramesh	Shah

Intersection

- As operações UNION e INTERSECTION são:
 - são comutativas, ou seja, $R \cup S = S \cup R$ e $R \cap S = S \cap R$
 - são aplicáveis a qualquer número de relações
 - são associativas, ou seja:
$$R \cup (S \cup T) = (R \cup S) \cup T$$
$$R \cap (S \cap T) = (R \cap S) \cap T$$

Operações da Teoria de Conjuntos

- Union
- Intersection
- **Difference**
- Produto Cartesiano
- Join



Difference

- Considerando as 2 relações união-compátíveis R e S, vistas anteriormente:
 - DIFFERENCE, denotada por $R - S$, tem como resultado uma relação que inclui todas as tuplas que estão em R, mas não estão em S

Difference

- Resultado da diferença de R - S

N	S
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

Difference

- Resultado da diferença de S - R

Nome	Sobrenome
John	Smith
Ricardo	Browne
Francis	Johnson

Difference

- A operação DIFFERENCE não é comutativa, pois em geral, $R-S \neq S-R$
- Note que a convenção adotada é que a relação resultante tem os mesmos nomes de atributos da primeira relação R

Operações da Teoria de Conjuntos

- Union
- Intersection
- Difference
- Produto Cartesiano
- Join



Produto Cartesiano

- Combina tuplas de duas relações R e S, resultando em uma relação que tem uma tupla para cada combinação de tuplas - uma de R e outra de S:

$$Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m) \leftarrow R(A_1, A_2, \dots, A_n) \times S(B_1, B_2, \dots, B_m)$$

Produto Cartesiano

- Exemplo: considere as seguintes relações R1 e R2

R1

NOME	SOBREN
João	Silva
Francisco	Souza

R2

NROD	NOMED
1	Pesquisa
2	Administração

Produto Cartesiano

- Exemplo do produto cartesiano de R1 e R2

R1XR2

NOME	SOBREN	NROD	NOMED
João	Silva	1	Pesquisa
João	Silva	2	Administração
Francisco	Souza	1	Pesquisa
Francisco	Souza	2	Administração

Produto Cartesiano

- Utilizando a **Linguagem SQL**, o produto cartesiano de R1 e R2, seria:

```
SELECT *
```

```
FROM R1, R2;
```


Operações da Teoria de Conjuntos

- Union
- Intersection
- Difference
- Produto Cartesiano
- Join



Join

- A operação Join, denotada por $|X|$, é utilizada para combinar tuplas de duas relações, através de um ou mais atributos comuns às duas relações
- A tabela resultante contém as colunas das duas tabelas que participaram da junção:

$$Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m) \leftarrow R(A_1, A_2, \dots, A_n) |X| S(B_1, B_2, \dots, B_m)$$

Join

- Esta operação é muito importante para qualquer banco de dados relacional, pois permite processar relacionamentos entre relações
- A forma geral de uma operação JOIN em 2 relações $R(A_1, A_2, \dots, A_n)$ e $S(B_1, B_2, \dots, B_m)$ é:

$R \mid X \mid S$ <condição de junção>

Join

- Considere as seguintes relações para o exemplo de join:

Empregado = {NOME, SOBREN, RG, NROD}

- NROD é chave estrangeira que referencia o atributo NDEPTO da relação Departamento

Departamento = {NDEPTO, NOMED, RGGERENTE }

- RGGERENTE é chave estrangeira que referencia o atributo RG da relação Empregado

Empregado

NOME	SOBREN	<u>RG</u>	NROD
João	Silva	123456789	5
Francisco	Souza	333445555	5
Alice	Fernandes	999887777	4
Joana	Pereira	987654321	4
Rodolfo	Nogueira	666884444	5

Departamento

NDEPTO	NOMED	RGGERENTE
1	Pesquisa	333445555
2	Administração	987654321
3	Finanças	666884444

Join

- Suponha que se queira recuperar o nome de cada departamento e o nome e sobrenome do gerente de cada departamento
- Para obter o nome do gerente, é necessário combinar cada tupla de Departamento com a tupla de Empregado cujo valor de RG seja igual ao valor de RGGERENTE na tupla de Departamento

$\text{DEPT_GER} \leftarrow \text{DEPARTAMENTO} \bowtie_{\text{RGGERENTE}=\text{RG}} \text{EMPREGADO}$

$\text{RESULT} \leftarrow \pi_{\text{NOMED, NOME, SOBREN}}(\text{DEPT_GER})$

Join

- Resultado da relação DEPT_GER

NDEPTO	NOMED	RGGERENTE	NOME	SOBREN	RG	NROD
1	Pesquisa	333445555	Francisco	Souza	333445555	5
2	Administração	987654321	Joana	Pereira	987654321	4
3	Finanças	666884444	Rodolfo	Nogueira	666884444	5

Join

- Resultado da relação RESULT

NOMED	NOME	SOBREN
Pesquisa	Francisco	Souza
Administração	Joana	Pereira
Finanças	Rodolfo	Nogueira

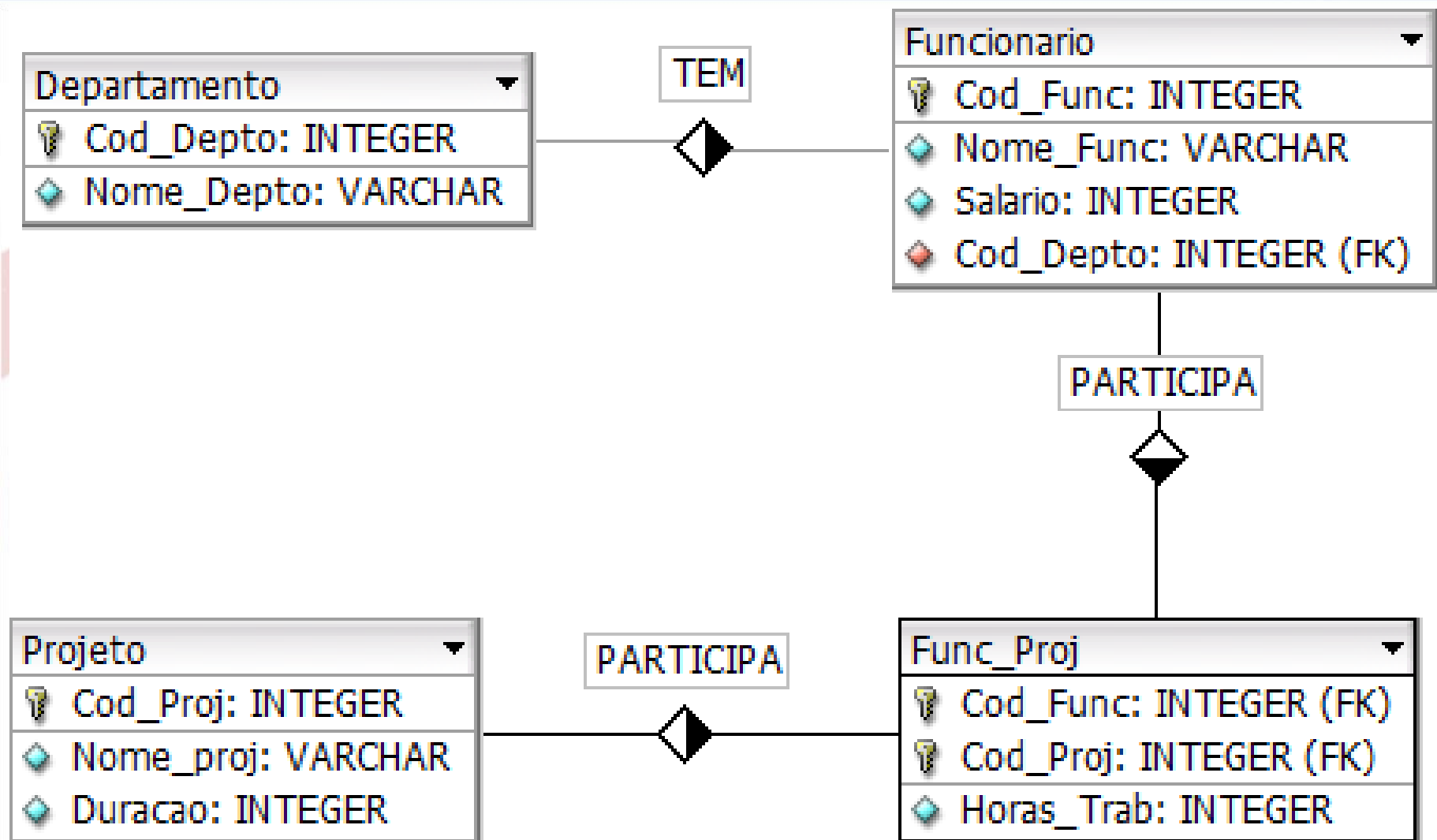


EXERCÍCIOS DE ÁLGEBRA RELACIONAL

Exercícios de Álgebra Relacional

Considere o seguinte Banco de Dados para esta aula (chave primária está sublinhada)

- Departamento = {Cod_Depto, Nome_Depto}
- Funcionario = {Cod_Func, Nome_Func, Salario, Cod_Depto}
 - Cod_depto é chave estrangeira que referencia o atributo Cod_depto da tabela Departamento
- Projeto = {Cod_Proj, Nome_Proj, Duracao}
- Func_Proj = {Cod_Func, Cod_Proj, Horas_Trab}
 - Cod_Func é chave estrangeira que referencia o atributo Cod_Func da tabela Funcionario
 - Cod_Proj é chave estrangeira que referencia o atributo Cod_Proj da tabela Projeto



Departamento

Cod_Depto	Nome_Depto
1	Marketing
2	Vendas
3	Dados
4	Pesquisa

Funcionário

Cod_Func	Nome_Func	Salario	Cod_Depto
101	Joao da Silva	2000	2
102	Mario Souza	1500	1
103	Sergio Santos	2400	2
104	Maria Castro	1200	1
105	Marcio Santana	1400	4

Projeto

Cod_Proj	Nome_Proj	Duracao
1001	Sistema A	2
1002	Sistema B	6
1003	Sistema X	4

Func_Proj

Cod_Func	Cod_Proj	Horas_Trab
101	1001	24
101	1002	160
102	1001	56
102	1003	45
103	1001	86
103	1003	64
104	1001	46

Exercícios de Álgebra Relacional

- 1) Selecione todos os dados dos empregados que ganham mais que 1600.
- 2) Selecione todos os dados dos empregados que ganham menos que 3000 e que pertencem ao departamento de código 2.
- 3) Selecione somente o nome de todos os departamentos.
- 4) Selecione o nome e duração dos projetos.

Exercícios de Álgebra Relacional

- 5) Selecione o nome e salário dos empregados que são do departamento de código 2.
- 6) Selecione o nome e duração dos projetos com mais de 3 meses de duração.
- 7) Selecione o código dos projetos e o número de horas trabalhadas nestes projetos, mas apenas para o empregado de código 102.

Exercícios de Álgebra Relacional

- 8) Selecione o nome dos empregados que são do departamento de código 2 e ganham um valor maior ou igual a 2000, ou então, são do departamento de código 1 e ganham mais de 1000.
- 9) Obtenha o nome do departamento que o funcionário 'Mário Souza' trabalha.
- 10) Obtenha o nome dos funcionários que trabalham no projeto de código 1001.

Exercícios de Álgebra Relacional

- 11)** Selecione o nome dos funcionários e o nome dos projetos que cada um trabalhou.
- 12)** Obtenha o nome dos projetos que a funcionária 'Maria Castro' trabalhou.
- 13)** Obtenha o nome e duração dos projetos onde houve a participação dos funcionários do departamento de 'Marketing'.

Resolução do exercício 12)

1ª forma – sem relações intermediárias

$\pi_{\text{Nome_Proj}}(\sigma_{(\text{Nome_Func} = \text{'Maria Castro'})}$
 $(\text{Funcionario} \bowtie_{\text{Cod_Func} = \text{Cod_Func}}$
 $(\text{Func_Proj} \bowtie_{\text{Cod_Proj} = \text{Cod_Proj}} \text{Projeto)))$

Resolução do exercício 12)

2ª forma – com relações intermediárias

Temp1 \leftarrow Func_Proj | X | $\text{Cod_Proj} = \text{Cod_Proj}$ Projeto

Temp2 \leftarrow Funcionario | X | $\text{Cod_Func} = \text{Cod_Func}$ Temp1

Temp3 $\leftarrow \sigma_{(\text{Nome_Func} = \text{'Maria Castro'})}$ (Temp2)

Result $\leftarrow \pi_{\text{Nome_Proj}}$ (Temp3)

Resolução do exercício 12)

3ª forma – com relações intermediárias e mais otimizada do que a 1ª e 2ª formas

$\text{Temp1} \leftarrow \sigma_{(\text{Nome_Func} = \text{'Maria Castro'})}(\text{Funcionario})$

$\text{Temp2} \leftarrow \text{Func_Proj} \bowtie_{\text{Cod_Func} = \text{Cod_Func}} \text{Temp1}$

$\text{Temp3} \leftarrow \text{Projeto} \bowtie_{\text{Cod_Proj} = \text{Cod_Proj}} \text{Temp2}$

$\text{Result} \leftarrow \pi_{\text{Nome_Proj}}(\text{Temp3})$

Obrigado

Profa. Elisângela Botelho Gracias
elisangela.botelho@mackenzie.br