

Banco de Dados I

06 - Álgebra Relacional

Arthur Porto - IFNMG Campus Salinas

arthur.porto@ifnmg.edu.br

arthurporto.com.br

Sumário I

- 1 Introdução
- 2 SELEÇÃO (σ)
- 3 PROJEÇÃO (π)
- 4 Sequência de operações
- 5 Exercício
- 6 Operações de conjuntos
 - UNIÃO (\cup)
 - INTERSECÇÃO (\cap)
 - DIFERENÇA ($S - I$)
 - DIFERENÇA ($I - S$)
 - PRODUTO CARTESIANO (\times)
- 7 JUNÇÃO (\bowtie)
 - EQUIJUNÇÃO e JUNÇÃO NATURAL
- 8 DIVISÃO \div
- 9 Extra

Sumário II

- Auto Relacionamento
- JUNÇÃO EXTERNA

10 Exemplo

11 Funções de Agregação

12 Referências

Sumário I

- 1 Introdução
- 2 SELEÇÃO (σ)
- 3 PROJEÇÃO (π)
- 4 Sequência de operações
- 5 Exercício
- 6 Operações de conjuntos
 - UNIÃO (\cup)
 - INTERSECÇÃO (\cap)
 - DIFERENÇA ($S - I$)
 - DIFERENÇA ($I - S$)
 - PRODUTO CARTESIANO (\times)
- 7 JUNÇÃO (\bowtie)
 - EQUIJUNÇÃO e JUNÇÃO NATURAL
- 8 DIVISÃO \div
- 9 Extra

Sumário II

- Auto Relacionamento
- JUNÇÃO EXTERNA

10 Exemplo

11 Funções de Agregação

12 Referências

- O conjunto básico de operações para o modelo relacional é a álgebra relacional
- As operações da álgebra produzem novas relações
- Características [1]
 - Alicerce formal para as operações do modelo relacional.
 - Base para implementação e otimização de consultas nos SGBDs
 - Alguns dos seus conceitos são incorporados à SQL
- Operações da teoria de conjuntos
 - União, intersecção, diferença e produto cartesiano
- Operações para BDs relacionais
 - Seleção, projeção e junção.

SELEÇÃO (σ)

- Usada para escolher um subconjunto das tuplas de uma relação que satisfaça uma condição de seleção.
- Representada pelo σ
- E a condição de seleção é uma expressão booleana.
 - A expressão pode conter várias condições usando os operadores:
 - **AND** (\wedge), **OR** (\vee) ou **NOT** (\neg)
 - Modelo da operação SELEÇÃO

$$\sigma_{condicao_de_selecao}(R) \quad (1)$$

- Exemplo

$$\sigma_{Dnr=4}(FUNCIONARIO) \quad (2)$$

SELEÇÃO (σ)

- Características

- A SELEÇÃO é unária
 - O operador SELEÇÃO é unário: aplica-se apenas a uma única relação (sem auto relacionamento também).
 - O grau da relação resultante é o mesmo da relação original.
 - O número de tuplas é menor ou igual a da relação original.
- A SELEÇÃO é comutativa
 - Uma sequência de SELEÇÃO pode ser aplicada em qualquer ordem.

$$\sigma_{\langle cond1 \rangle}(\sigma_{\langle cond2 \rangle}(R)) = \sigma_{\langle cond2 \rangle}(\sigma_{\langle cond1 \rangle}(R)) \quad (3)$$

$$\sigma_{\langle cond1 \rangle}(\sigma_{\langle cond2 \rangle}(R)) = \sigma_{\langle cond1 \rangle \text{ AND } \langle cond2 \rangle}(R) \quad (4)$$

PROJEÇÃO (π)

- A PROJEÇÃO seleciona colunas da tabelas (os atributos da relação).
- Representada pelo π
- É uma partição vertical da relação.
- Modelo da operação PROJEÇÃO

$$\pi_{lista_de_atributos}(R) \quad (5)$$

- Exemplo

$$\pi_{Unome,Pnome,Salario}(FUNCIONARIO) \quad (6)$$

PROJEÇÃO (π)

- Características

- O grau da relação resultante é igual o número e atributos na lista de atributos.
- Caso apenas atributos não-chave estejam na lista de atributos a PROJEÇÃO elimina quais quer tuplas duplicadas na relação resultante.
- O número de tuplas da relação resultante é sempre menor ou igual a da relação original.
- Não é comutativa!

Sequência de operações

- Exemplos

$$\text{FUNCS_DEPT5} = \sigma_{Dnr=5}(\text{FUNCIONARIO}) \quad (7)$$

- Utilização em outras operações

$$\text{TEMP} = \sigma_{Dnr=5}(\text{FUNCIONARIO}) \quad (8)$$

$$R = \pi_{Pnome, Unome, Salario}(\text{TEMP}) \quad (9)$$

- Relações
 - VIRUS(nomeCientifico, nomePopular, incubacao)
 - MEDICAMENTO(nomeVenda, compostoAtivo)
- Liste todos os vírus com período de incubação maior que 5 dias.
- Liste todos os componentes ativos disponíveis
- Liste o nome popular de todos os vírus com período de incubação maior que 5 dias.

- Relações
 - PESSOA(nome, nomeMae, anoNascimento, nomeCidadeNatal)
 - nomeCidadeNatal : CHE Cidade
 - CIDADE(nomeCidade, siglaEstado)
- Liste os nomes de todas as mães.
- Liste os nomes de todas as mães com filhos maiores de 12 anos.

Operações de conjuntos

- UNIÃO (\cup), INTERSECÇÃO (\cap) e DIFERENÇA ($-$)
- São usadas para mesclar elementos de dois conjuntos (operações binárias).
- É necessária a compatibilidade de união (tipo)
 - Duas relações $R(A_1, A_2, \dots, A_n)$ e $S(B_1, B_2, \dots, B_n)$ são compatíveis no tipo se tiverem o mesmo grau n e se $\text{dom}(A_i) = \text{dom}(B_i)$
 - Relações com o mesmo número de atributos e estes devem ter o mesmo domínio.
 - Exemplo: CPF de todos os funcionários que trabalham ou no Dep. 5 ou supervisionam funcionários no Dep. 5

$$\begin{aligned}FUNCS_DEP5 &= \sigma_{Dnr=5}(FUNCIONARIO) \\RES1 &= \pi_{Cpf}(FUNCS_DEP5) \\RES2 &= \pi_{Cpf_supervisor}(FUNCS_DEP5) \\RESFINAL &= RES1 \cup RES2\end{aligned}$$

Operações de conjuntos - UNIÃO (\cup)

STUDENT

F _n	L _n
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

INSTRUCTOR

Fname	Lname
John	Smith
Ricardo	Browne
Susan	Yao
Francis	Johnson
Ramesh	Shah

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

Operações de conjuntos - INTERSECÇÃO (\cap)

STUDENT

F _n	L _n
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

INSTRUCTOR

Fname	Lname
John	Smith
Ricardo	Browne
Susan	Yao
Francis	Johnson
Ramesh	Shah

F _n	L _n
Susan	Yao
Ramesh	Shah

Operações de conjuntos - DIFERENÇA ($S - I$)

STUDENT

F _n	L _n
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

INSTRUCTOR

Fname	Lname
John	Smith
Ricardo	Browne
Susan	Yao
Francis	Johnson
Ramesh	Shah

F _n	L _n
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

Operações de conjuntos - DIFERENÇA ($I - S$)

STUDENT

Fn	Ln
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

INSTRUCTOR

Fname	Lname
John	Smith
Ricardo	Browne
Susan	Yao
Francis	Johnson
Ramesh	Shah

Fname	Lname
John	Smith
Ricardo	Browne
Francis	Johnson

- Características

- UNIÃO e INTERSECÇÃO são comutativas e associativas

$$R \cup S = S \cup R \text{ e } R \cap S = S \cap R$$

$$R \cup (S \cup T) = (R \cup S) \cup T \text{ e } R \cap (S \cap T) = (R \cap S) \cap T$$

- DIFERENÇA não é comutativa

$$R - S \neq S - R$$

Operações de conjuntos - PRODUTO CARTESIANO (\times)

- Conhecida também como produto cruzado.
- Indicado pelo simbolo \times .
- Operação binária.
- O grau do produto cartesiano entre as relações $R(A_1, A_2, \dots, A_n)$ e $S(B_1, B_2, \dots, B_m)$ é $n + m$.
- Exemplo: Lista dos nomes dos dependentes de cada funcionária.

$$\begin{aligned} EMP_FEM &= \sigma_{Sexo='F'}(FUNCIONARIO) \\ EMPNOMES &= \pi_{Pnome, Unome, Cpf}(EMP_FEM) \\ EMP_DEP &= EMPNOMES \times DEPENDENTE \\ ATUAL_DEP &= \sigma_{Cpf=Fcpf}(EMP_DEP) \\ RESULT &= \pi_{Pnome, Unome, Nome.dependente}(ATUAL_DEP) \end{aligned} \tag{10}$$

JUNÇÃO (\bowtie)

- A JUNÇÃO é indicada por \bowtie .
- Usada para combinar tuplas relacionadas.
- É a composição de um produto cartesiano seguido de uma seleção.
- Exemplo: Nome do gerente de cada departamento.

$$\begin{aligned} DEP_GER &= DEPARTAMENTO \bowtie_{Cpf_{gerente}=Cpf} FUNCIONARIO \\ RESULTAO &= \pi_{Dnome, Unome, Pnome}(DEP_GER) \end{aligned} \quad (11)$$

- A JUNÇÃO pode ser realizada em múltiplas tabelas também.

$$(PROJETO \bowtie_{Dnum=Dnumero} DEPART) \bowtie_{Cpf_{gerente}=Cpf} FUNC \quad (12)$$

JUNÇÃO (\bowtie) - EQUIJUNÇÃO e JUNÇÃO NATURAL

- Relações

cliente			corrida		
<u>Clild</u>	Nome	CPF	<u>Clild</u>	<u>Placa</u>	<u>DataPedido</u>

- EQUIJUNÇÃO

$$CLIENTE \bowtie_{Clild} CORRIDA \quad (13)$$

- JUNÇÃO NATURAL

$$CLIENTE * CORRIDA \quad (14)$$

DIVISÃO ÷

- Na DIVISÃO

$$T = R(Z) \div S(X) \quad (15)$$

onde:

Z : subconjunto de X ($X \subseteq Z$)

- Para uma tupla t aparecer no resultado da DIVISÃO os valores em t deverão aparecer em R em combinação com cada tupla em S

R

A	B
a1	b1
a2	b1
a3	b1
a4	b1
a1	b2
a3	b2
a2	b3
a3	b3
a4	b3
a1	b4
a2	b4
a3	b4

S

A
a1
a2
a3

T

B
b1
b4

- Exemplo: Liste todos os funcionários que trabalham em todos os projetos em que 'João Silva' trabalha.

$$\begin{aligned}
 SILVA &= \sigma_{Pnome='Joao' \text{ AND } Unome='Silva'}(FUNC) \\
 SILVA_PNR &= \pi_{Pnr}(TRAB_EM \bowtie_{Fcpf=Cpf} SILVA) \\
 CPF_PNR &= \pi_{Fcpf, Pnr}(TRAB_EM) \\
 CPFS(Cpf) &= CPF_PNR \div SILVA_PNR \\
 RESULTADO &= \pi_{Pnome, Unome}(CPFS * FUNC)
 \end{aligned}$$

(16)

CPF_PNR	
Fcpf	Pnr
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

SILVA_PNR
Pnr
1
2

CPFS
Cpf
123456789
453453453

- Exemplo: Nome do funcionário e o nome do seu supervisor.

$FUNC = FUNCIONARIO$

$SUPER(F_Pnome, S_Cpf) = \pi_{Pnome, Cpf_supervisor}(FUNC)$ (17)

$RESULT = \pi_{F_Pnome, Pnome}(SUPER \bowtie_{S_Cpf=Cpf} FUNC)$

- JUNÇÃO EXTERNA À ESQUERDA ($R \bowtie S$)
- JUNÇÃO EXTERNA À DIREITA ($A \bowtie B$)
- JUNÇÃO EXTERNA COMPLETA ($A \bowtie B$)

Exemplo

- Exemplo com um pouco de tudo.

- Números de projeto com funcionários cujo último nome é 'Silva', seja como participante ou gerente do departamento que controla o projeto.

$$\begin{aligned} FUNC &= FUNCIONARIO \\ DEP &= DEPARTAMENTO \\ SILVA(Fcpf) &= \pi_{Cpf}(\sigma_{Unome=Silva}(FUNC)) \\ P_SILVA_TRAB &= \pi_{Pnr}(TRABALHA_EM * SILVA) \\ GER &= \pi_{Unome,Dnumero}(FUNC \bowtie_{Cpf=Cpf-gerente} DEP) \\ DEP_SIL_GER(Dnum) &= \pi_{Dnumero}(\sigma_{Unome=Silva}(GER)) \\ P_SIL_GER(Pnr) &= \pi_{Projnumero}(DEP_SIL_GER * PROJETO) \\ RES &= (P_SIL_TRAB \cup P_SIL_GER) \end{aligned} \tag{18}$$

Funções de Agregação

- As funções de agregação pegam uma coleção de valores e retorna um único valor como resultado [2].
- Símbolos: \mathfrak{G} ou \mathcal{G}
- Exemplos:
 - \mathcal{G}_{sum} , \mathcal{G}_{avg} , $\mathcal{G}_{\text{count}}$, \mathcal{G}_{min} e \mathcal{G}_{max}
 - Encontrar a soma dos salários de todos os professores:

$$\mathcal{G}_{\text{sum}(\text{salary})}(\text{professor}) \quad (19)$$

- Quando é necessário eliminar múltiplas ocorrências use o **distinct**.
 - Encontre o número total de professores que lecionam uma disciplina no segundo semestre de 2010.

$$\mathcal{G}_{\text{count}-\text{distinct}(ID)}(\sigma_{\text{professor}=\text{"segundo"} \wedge \text{ano}=2010}(\text{ensina})) \quad (20)$$

Funções de Agregação

- Quando é necessário aplicar uma função agregada a grupos de tuplas.
 - Encontre o salário médio de cada departamento.

$$\text{dept_name } \mathcal{G}_{\text{avg}(\text{salary})}(\text{instructor}) \quad (21)$$

ID	name	dept_name	salary
76766	Crick	Biology	72000
45565	Katz	Comp. Sci.	75000
10101	Srinivasan	Comp. Sci.	65000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000
12121	Wu	Finance	90000
76543	Singh	Finance	80000
32343	El Said	History	60000
58583	Califieri	History	62000
15151	Mozart	Music	40000
33456	Gold	Physics	87000
22222	Einstein	Physics	95000

dept_name	salary
Biology	72000
Comp. Sci.	77333
Elec. Eng.	80000
Finance	85000
History	61000
Music	40000
Physics	91000

Referências



R. Elmasri and S.B. Navathe.
Sistemas de banco de dados.
PEARSON BRASIL, 2011.



A. Silberschatz, S. Sundarshan, and H.F. Korth.
Sistema de banco de dados.
Elsevier, 2011.