



1948 - 2008

Álgebra relacional – revisão de alguns conceitos

Galina

Conceitos necessários para otimização de
consultas

Slides adaptados do material do prof Heuser
(UFRGS)

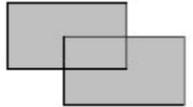
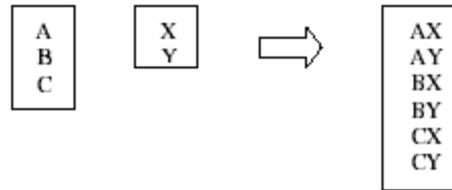
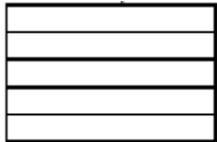
Álgebra relacional

- Álgebra desenvolvida para descrever **operações** sobre uma base de dados **relacional**
- Os objetos **sobre** os quais a álgebra opera são **tabelas**
- Uma operação possui como **operandos** e como resultado **tabelas**
- Por que aprender:
 - Compreendendo álgebra relacional é mais fácil apreender **SQL**
 - Não há SGBD que implementa álgebra diretamente como DML, mas SQL **incorpora** cada vez mais conceitos de álgebra
 - **Algoritmos de otimização de consulta definidos sobre álgebra (uso internamente no SGBD)**

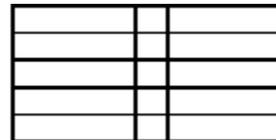
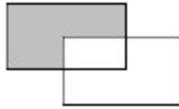
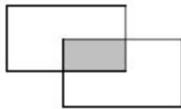
Operadores da álgebra relacional

- **Operadores originários da teoria de conjuntos (uma tabela é um conjunto de linhas):**
 - União
 - Interseção
 - Diferença
 - Produto Cartesiano
- **Operadores *específicos* da álgebra relacional:**
 - Seleção
 - Projeção
 - Junção
 - Divisão
 - Renomeação

Representação Gráfica



Junção



Simbologia

- União $R \cup S$
- Intersecção $R \cap S$
- Diferença $R - S$
- Prod. Cartesiano $R \times S$
- Seleção $\sigma_F (R)$
- Projeção $\pi_{i\theta_j^{i_1, i_2, \dots, i_m}} (R)$
- Junção $R \bowtie S$

BD Exemplo

Peça

CodPeça	NomePeça	CorPeça	PesoPeça	CidadePeça
P1	Eixo	Cinza	10	PoA
P2	Rolamento	Preto	16	Rio
P3	Mancal	Verde	30	SãoPaulo

Embarq

CodPeça	CodFornec	QtdeEmbarc
P1	F1	300
P1	F2	400
P1	F3	200
P2	F1	300
P2	F4	350

Fornec

CodFornec	NomeFornec	StatusFornec	CidadeFornec
F1	Silva	5	SãoPaulo
F2	Souza	10	Rio
F3	Álvares	5	SãoPaulo
F4	Tavares	8	Rio

Operador de seleção

- A **Seleção** tem como operando uma tabela.
- O resultado é uma tabela que contém as linhas que obedecem a um determinado **critério**.
- **Sintaxe:**
 $\sigma\langle\text{critério de seleção}\rangle (\langle\text{tabela}\rangle)$
 - Onde:
 - $\langle\text{tabela}\rangle$ é o nome de uma **tabela** ou uma **expressão** de álgebra relacional que resulta em uma tabela
 - $\langle\text{critério de seleção}\rangle$ é uma **expressão** booleana que envolve literais e **valores** de atributos da tabela
- O resultado da seleção tem colunas com os mesmos nomes e domínios da tabela de entrada.

Exemplo de Seleção (BD de peças e fornecedores)

σ CodPeça = 'P1' (Peça)

- Resulta em uma tabela que contém todos os dados da peça de código P1

CodPeça	NomePeça	CorPeça	PesoPeça	CidadePeça
P1	Eixo	Cinza	10	PoA

Exemplo de Seleção

σ (StatusFornec > 5 and CidadeFornec = 'Rio')
(Fornec)

- Resulta em uma tabela com os dados de todas os fornecedores que tenham status maior que 5 e sejam do Rio.

CodFornec	NomeFornec	StatusFornec	CidadeFornec
F2	Souza	10	Rio
F4	Tavares	8	Rio

Seleção é comutativa

σ (StatusFornec > 5) (σ (CidadeFornec = 'Rio') (Fornec))

- O mesmo que o anterior, mas usando duas expressões de Seleção *aninhadas*
- O critério de seleção envolve somente valores de atributos de *uma linha*. Não é possível especificar critérios que envolvam múltiplas linhas nem critérios que envolvem diferentes tabelas.

Operação de Projeção

- A **Projeção** tem como operando uma tabela. O resultado é uma tabela que contém apenas as colunas selecionadas.
- Sintaxe:

π <lista de colunas> (<tabela>)

– Onde:

- <tabela> é o nome de uma **tabela** ou uma **expressão** de álgebra relacional que resulta em uma tabela
- <lista de **colunas**> é uma lista que contém nomes de **colunas** da tabela operando.

Exemplo de Projeção (BD de peças e fornecedores)

π CodPeça, NomePeça (Peça)

- Resulta em uma tabela que contém os códigos e os nomes de todas as peças.

CodPeça	NomePeça
P1	Eixo
P2	Rolamento
P3	Mancal

Aninhamento de operadores

- Operadores diferentes podem ser aninhados

π CodFornec, QtdeEmbarc (σ CodPeça = 'P1'
(Embarq))

- Resulta em tabela com código de fornecedor e quantidade embarcada para cada embarque da peça com código 'P1'.

CodFornec	QtdeEmbarc
F1	300
F2	400
F3	200

Operações da teoria de conjuntos

- A álgebra relacional empresta da teoria de conjuntos quatro operadores: União, Intersecção, Diferença e Produto Cartesiano
- Sintaxe da operação **União**: $\langle \text{tabela} \rangle_1 \cup \langle \text{tabela} \rangle_2$
- Sintaxe da operação **Intersecção**: $\langle \text{tabela} \rangle_1 \cap \langle \text{tabela} \rangle_2$
- Sintaxe da operação **Diferença**: $\langle \text{tabela} \rangle_1 - \langle \text{tabela} \rangle_2$
- Nos três casos, a operação possui duas tabelas como operando. As tabelas devem ser **compatíveis** para união:
 - Possuir o mesmo **número de colunas**
 - O **domínio** da *i-ésima* coluna de uma tabela deve ser idêntico ao domínio da *i-ésima* coluna da outra.
- Quando os nomes das colunas forem diferentes, adota-se a convenção de usar os nomes das colunas da primeira tabela.

Operação Produto Cartesiano

- Sintaxe da operação Produto Cartesiano:

$$\langle \text{tabela} \rangle_1 \mathbf{X} \langle \text{tabela} \rangle_2$$

- O produto cartesiano possui como operandos **duas tabelas**.
- O resultado é uma tabela cujas linhas são a combinação das linhas das tabelas $\langle \text{tabela} \rangle_1$ e $\langle \text{tabela} \rangle_2$, tomando-se uma linha da $\langle \text{tabela} \rangle_1$ e concatenando-a com uma linha da $\langle \text{tabela} \rangle_2$.
- Total de colunas do produto cartesiano = Número de colunas da primeira tabela + Número de colunas da segunda tabela
- Número de linhas do produto cartesiano = Número de linhas da primeira tabela x Número de linhas da segunda tabela.

Exemplo de Produto Cartesiano

- Embarq

Embarq			Peça				
CodPeça	CodFornec	QtdeEmbarc	CodPeça	NomePeça	CorPeça	PesoPeça	CidadePeça
P1	F1	300	P1	Eixo	Cinza	10	PoA
P1	F2	400	P1	Eixo	Cinza	10	PoA
P1	F3	200	P1	Eixo	Cinza	10	PoA
P2	F1	300	P1	Eixo	Cinza	10	PoA
P2	F4	350	P1	Eixo	Cinza	10	PoA
P1	F1	300	P2	Rolamento	Preto	16	Rio
P1	F2	400	P2	Rolamento	Preto	16	Rio
P1	F3	200	P2	Rolamento	Preto	16	Rio
P2	F1	300	P2	Rolamento	Preto	16	Rio
P2	F4	350	P2	Rolamento	Preto	16	Rio
P1	F1	300	P3	Mancal	Verde	30	SãoPaulo
P1	F2	400	P3	Mancal	Verde	30	SãoPaulo
P1	F3	200	P3	Mancal	Verde	30	SãoPaulo
P2	F1	300	P3	Mancal	Verde	30	SãoPaulo
P2	F4	350	P3	Mancal	Verde	30	SãoPaulo

- A operação Produto Cartesiano não é usada isoladamente.
 - Normalmente, ela é combinada com uma seleção que envolve as diversas tabelas multiplicadas.

Seleção combinada com Produto Cartesiano

π NomePeça (σ Embarq.CodPeça=Peça.CodPeça
(Embarq X Peça))

- Obtém os nomes de todas as peças para as quais há embarques

NomePeça
Eixo
Rolamento

Outros operadores

- União: $A \cup B$
 - o conjunto de todas as tuplas pertencentes a relação A ou pertencentes a relação B
- Diferença: $A - B$
 - o conjunto de todas as tuplas pertencentes a relação A e não pertencentes a relação B.
- Interseção: $A \cap B$
 - O conjunto de todas as tuplas pertencentes a relação A e pertencentes a relação B.

Operadores derivados

- Observar que há operadores de álgebra que são deriváveis de outros.
- Exemplo:
 - Operação de interseção é derivável de união e diferença
 - $A \cap B = A - (A - B)$

Operação de Junção

- A combinação de uma operação de **seleção** aplicada sobre uma operação de **produto cartesiano** é usual em aplicações de BD. É através dela que **dados** de tabelas relacionadas **são associados**. Por isso, foi criada a operação de junção que corresponde exatamente à seqüência de operações em questão.

- **Sintaxe:**

– Onde: $\langle \text{tabela} \rangle_1 \bowtie \langle \text{critério} \rangle \langle \text{tabela} \rangle_2$

- $\langle \text{tabela} \rangle$ é o nome de uma **tabela** ou uma **expressão** de álgebra relacional que resulta em uma tabela
 - $\langle \text{critério} \rangle$ é uma **expressão** booleana envolvendo literais e valores de atributos das duas tabelas.
- A Junção tem como operandos duas tabela. O resultado é equivalente a executar:

$\sigma \langle \text{critério} \rangle (\langle \text{tabela} \rangle_1 \times \langle \text{tabela} \rangle_2)$

Exemplo de Junção (BD de peças e fornecedores)

Embarq ⋈ (Embarq.CódFornec= Fornec.CódFornec) Fornec

- Associa cada linha de embarque com a correspondente linha de fornecedor.

Embarq			Fornec			
CodPeça	CodFornec	QtdeEmbarc	CodFornec	NomeFornec	StatusFornec	CidadeFornec
P1	F1	300	F1	Silva	5	SãoPaulo
P1	F2	400	F2	Souza	10	Rio
P1	F3	200	F3	Álvares	5	SãoPaulo
P2	F1	300	F1	Silva	5	SãoPaulo
P2	F4	350	F4	Tavares	8	Rio

Junção theta, Equijunção e Junção natural

- Critério de junção:
 - qualquer expressão booleana, inclusive comparações do tipo $<$, $>$, $<>$, ... entre os valores de atributos das tabelas envolvidas na junção
 - Essa operação **genérica** de Junção é chamada de *Junção theta*.
- Na **maior parte dos casos**, o $<$ critério $>$ de junção é uma expressão como mostrada no exemplo
 - Envolve apenas **igualdade** de valores de atributos de diferentes tabelas
 - Esse tipo de junção é chamada de *Equijunção*

Exemplo de Equijunção (BD de peças e fornecedores)

Embarq \bowtie (CódFornec), (CódFornec) Fornec

- Associa cada linha de embarque com a correspondente linha de fornecedor.

Embarq			Fornec		
CodPeça	CodFornec	QtdeEmbarc	NomeFornec	StatusFornec	CidadeFornec
P1	F1	300	Silva	5	SãoPaulo
P1	F2	400	Souza	10	Rio
P1	F3	200	Álvares	5	SãoPaulo
P2	F1	300	Silva	5	SãoPaulo
P2	F4	350	Tavares	8	Rio

Junção Natural

- No caso acima, as colunas de junção possuem os **mesmos nomes**.
- Para estes casos existe a *Junção natural*, na qual as listas de nomes de colunas não necessitam ser especificadas.
- Sintaxe da Junção natural:

`<tabela>1 ⋈ <tabela>2`

- Exemplo de Junção natural
 - Associa cada linha de embarque com a correspondente linha de fornecedor.

`Embarq ⋈ Fornec`

Junção externa (outer join)

- Exemplo:

Empregado  (CódEmp=CodEmpGer) Departamento

- O operador  e chamado de *Junção externa esquerda* (“*left outer-join*”)

- Semântica

- A Junção externa esquerda contém ao menos uma vez cada linha da tabela à esquerda do operador (no caso a tabela Empregado). Esta linha aparece concatenada com uma linha vazia, caso o critério de junção não seja verdadeiro para nenhuma linha da tabela à direita do operador de junção. Caso o critério de junção seja verdadeiro para uma ou mais linhas da tabela à direita, a linha da tabela à esquerda aparecerá concatenada com uma ou mais linhas da tabela à direita.

- De forma similar podem ser definidas:

- Junção externa direita - símbolo
- Junção externa plena - símbolo





Conjunto mínimo de operações

- Muitas operações podem ser derivadas de outras
- Foi identificado um conjunto mínimo (completo) de operações, das quais todas as demais operações da álgebra original podem ser derivadas:
 - Seleção
 - Projeção
 - União
 - Diferença
 - Produto Cartesiano