

Álgebra Relacional

Bancos de Dados I

Altigran Soares da Silva

IComp/UFAM - 2018/02

Adaptado do Material do Professor Jeffrey Ullman

O que é uma “Álgebra”

- Sistema matemático que consiste de:
 - **Operandos**: variáveis ou valores a partir dos quais novos valores podem ser gerados
 - **Operadores** --- símbolos que denotam operações que geram novos valores a partir de valores existentes

2

O que é Álgebra Relacional?

- Um Álgebra cujos operandos são relações ou variáveis que representam relações
- Os operandos foram projetados de forma a cobrir operações comuns sobre relações em um Banco de dados

3

Por que Álgebra Relacional?

- Não é usada hoje como linguagem de consulta em SGBDs comerciais
 - Embora alguns dos primeiros protótipos tenham usado.
- Em vez disso, a linguagem de consulta "real", a SQL, incorpora a AR em seu cerne
 - Muitos comandos em SQL são na realidade expressões sintaticamente "açucaradas" da AR
- No processamento de consultas em um SGBD, a primeira coisa que acontece a uma consulta SQL é sua tradução em AR ou em uma representação interna muito semelhante.

4

Operações Básicas

■ União, Intersecção e Diferença

- ❑ Operações usuais sobre conjuntos, mas os dois operandos devem ter um mesmo esquema

■ Seleção: filtra um subconjunto das linhas

■ Projeção: filtra um subconjunto das colunas

■ Produtos e Junções: combina dados de tabelas diferentes

■ Renomeação de relações e atributos

5

Seleção

■ $R1 := \sigma_C(R2)$

- ❑ C é uma condição que se refere aos atributos de $R2$.
- ❑ $R1$ é formada pelas tuplas de $R2$ que satisfazem C .
- ❑ $R1$ tem o mesmo esquema de $R2$

6

Exemplo: Seleção

Tabela Vendas

bar	cerveja	valor
Zeca's	Bud	2,50
Zeca's	Miller	2,75
Coroado	Bud	2,50
Coroado	Miller	3,00

ZecaMenu := $\sigma_{\text{bar}=\text{"Zeca's"}}(\text{Vendas})$:

bar	cerveja	valor
Zeca's	Bud	2,50
Zeca's	Miller	2,75

7

Projeção

■ $R1 := \pi_L(R2)$

- ❑ L é uma lista de atributos de $R2$.
- ❑ $R1$ é construída tomando cada tupla de $R2$, e extraíndo somente os atributos da lista L , na ordem especificada
- ❑ Se houverem tuplas duplicadas, estas são eliminadas

8

Exemplo: Projeção

Tabela Vendas

bar	cerveja	valor
Zeca's	Bud	2,50
Zeca's	Miller	2,75
Coroado	Bud	2,50
Coroado	Miller	3,00

valores := $\pi_{\text{cerveja, valor}}(\text{Vendas})$:

cerveja	valor
Bud	2.50
Miller	2.75
Miller	3.00

9

Projeção Extendida

- Usando o mesmo operador π_L , a lista L pode conter expressões arbitrárias sobre os atributos:
 1. Aritmética sobre os atributos
 - Ex: $A+B \rightarrow C$.
 2. Vária ocorrências do mesmo atributo

10

Exemplo: Projeção Estendida

$R =$

A	B
1	2
3	4

$\pi_{A+B \rightarrow C, A, A}(R) =$

C	A1	A2
3	1	1
7	3	3

11

Produto Cartesiano

- $R3 := R1 \times R2$
 - Emparelha cada tupla $t1$ de $R1$ com cada tupla $t2$ de $R2$.
 - Cada tupla de $R3$ é um concatenação $t1t2$
 - O esquema de $R3$ é composto por todos os atributos de $R1$ e $R2$, na ordem.
 - Se houver um atributo com o mesmo nome A em $R1$ e $R2$: usar $R1.A$ e $R2.A$.

12

Exemplo: $R3 := R1 \bowtie R2$

R1(A, B)

1	2
3	4

R2(B, C)

5	6
7	8
9	10

R3(A, R1.B, R2.B, C)

1	2	5	6
1	2	7	8
1	2	9	10
3	4	5	6
3	4	7	8
3	4	9	10

13

Junção Theta

- $R3 := R1 \bowtie_C R2$
 - Toma o $R1 \times R2$.
 - Aplica σ_C ao resulta
- Em σ , C pode ser qualquer condição booleana
 - Versões históricos só permitiam operadores $\theta =, <, \text{etc.}$; por isso o nome “junção theta”

14

Exemplo: Junção Theta

Vendas

bar	cerveja	valor
Zeca's	Bud	2.50
Zeca's	Miller	2.75
Coroado	Bud	2.50
Coroado	Coors	3.00

Bares

nome	endereço
Zeca's	R.Octávio
Coroado	Beira Rio

BarInfo := Vendas $\bowtie_{\text{Vendas.bar} = \text{Bares.nome}}$ Bares

BarInfo

bar	cerveja	valor	nome	endereço
Zeca's	Bud	2.50	Zeca's	R. Octávio
Zeca's	Miller	2.75	Zeca's	R. Octávio
Coroado	Bud	2.50	Coroado	Beira Rio
Coroado	Coors	3.00	Coroado	Beira Rio

Junção Natural

- Junção com atributos de mesmo nome
 - Aplica comparação de igualdade entre atributos de mesmo nome
 - Somente um dos atributos de mesmo nome é mantido.
- Denotada por $R3 := R1 \bowtie R2$.

16

Exemplo: Junção Natural

Vendas

bar	cerveja	valor
Zeca's	Bud	2.50
Zeca's	Miller	2.75
Coroado	Bud	2.50
Coroado	Coors	3.00

Bares

bar	endereço
Zeca's	R.Octávio
Coroado	Beira Rio

BarInfo := Vendas \bowtie Bares

BarInfo

bar	cerveja	valor	endereço
Zeca's	Bud	2.50	R. Octávio
Zeca's	Miller	2.75	R. Octávio
Coroado	Bud	2.50	Beira Rio
Coroado	Coors	3.00	Beira Rio

Renomeação

- Operador ρ
 - modifica o esquema da relação
- $R1 := \rho_{R1(A1,...,An)}(R2)$
 - cria R1 com os atributos $A1,...,An$
 - R1 tem as mesmas tuplas de R2.
- Notação Simplificada: $R1(A1,...,An) := R2$.

18

Exemplo: Renomeação

Bares

nome	endereço
Zeca's	R. Octávio
Coroado	Beira Rio

Onde(bar, local) := Bares

Onde

bar	local
Zeca's	R. Octávio
Coroado	Beira Rio

Construindo Expressões Complexas

- Operadores podem ser combinados
 - Regras Precedência e Parênteses
- Três notações, como em aritmética
 1. Sequências de comandos de atribuição
 2. Expressões com múltiplos operadores.
 3. Árvores de Expressão

Sequências de Atribuições

- Utiliza nomes de relações temporárias
- Pode ocorrer renomeação implícita de atributos
- **Exemplo:** $R3 := R1 \bowtie_c R2$ pode ser escrita:
 $R4 := R1 \times R2$
 $R3 := \sigma_c(R4)$

21

Expressões com Múltiplos Operadores

- Exemplo: $R3 := R1 \bowtie_c R2$
- $R3 := \sigma_c(R1 \times R2)$
- Precedência de Operadores:
 - $[\sigma, \pi, \rho]$ (maior).
 - $[X, \bowtie]$.
 - \cap .
 - $[U, -]$

22

Árvores de Expressão

- Folhas são operandos
 - Variables representando relações ou as próprias relações
- Nós internos são operadores aplicados aos nó(s) filho(s)

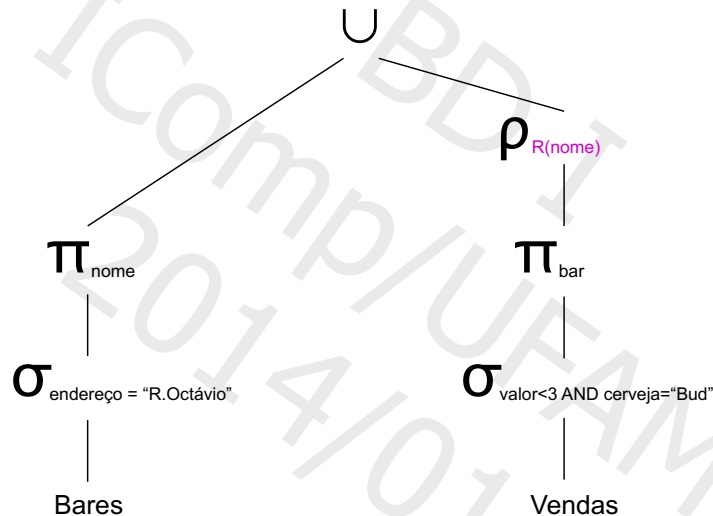
23

Exemplo: Árvores de Expressão

- Sejam as relações **Bares(nome, endereço)** e **Vendas(bar, cerveja, valor)**
- Defina uma expressão para encontrar os nomes dos bares que se encontram na “R.Octávio” ou que vendem Bud por menos de R\$3,00.

24

Árvore de Expressão



25

Exemplo: Auto-Junção

- Considere **Vendas(bar, cerveja, valor)**
- Quais os bares que vendem duas cervejas diferentes pelo mesmo preço?

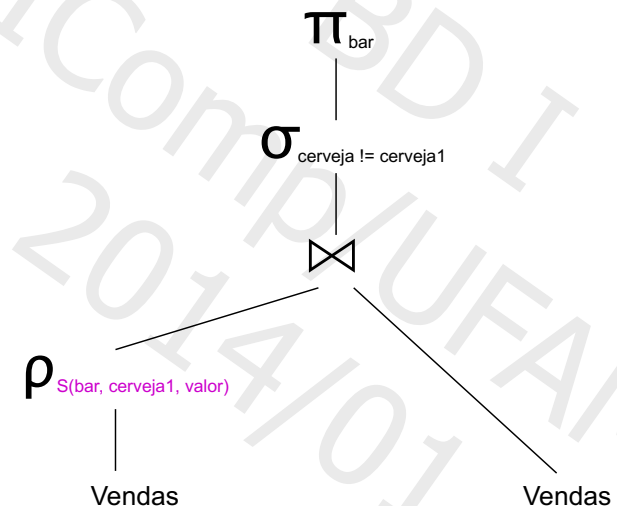
26

Exemplo: Auto-Junção

- Estratégia:
 - Defina, usando renomeação, uma cópia de **Vendas** da seguinte forma **S(bar, cerveja1, valor)**
 - A junção natural de **Vendas** com **S** contém quádruplas **(bar, cerveja, cerveja1, valor)** tal que o bar vende as duas cervejas pelo mesmo preço

27

Árvore de Expressão



28

Esquema dos Resultados

- União, interseção e diferença:
 - Os esquemas dos dois operandos devem ser os mesmos. Esse esquema será o do resultado.
- Seleção:
 - esquema do resultado é o mesmo que o esquema do operando.
- Projeção:
 - lista de atributos define o esquema

29

Esquema dos Resultados

- Produto e Junção-Theta:
 - esquema contém todos os atributos de ambas as relações
 - Atributos com o mesmo nome podem ser usados com prefixos, ex., R.A
- Junção natural:
 - união dos atributos das duas relações.
- Renomear:
 - o operador determina o esquema

30

Álgebra Relacional sobre Bags

- **Bags** (ou multi-conjuntos) são conjuntos onde um mesmo elemento pode ocorrer mais de uma vez
 - $\{1,2,1,3\}$ é um bag mas não é um conjunto.
 - $\{1,2,3\}$ é um bag que também é um conjunto.

31

Por que Bags?

- A SQL, linguagem mais importante para BDs relacionais, atua na realidade sobre Bags.
- Alguns operadores, como a projeção, são mais efetivos sobre bags que sobre conjuntos.

32

Operações sobre Bags

- **Seleção**: se aplica a cada tupla, por isso seu efeito sobre bags é o mesmo que sobre conjuntos.
- **Projeções** quando aplicada sobre bags não elimina duplicatas.
- **Produtos** e **Junções** são aplicadas sobre pares de tuplas. Portanto duplicatas em bags não trazem efeitos.

33

Exemplo: Seleção em Bags

$R($

A,	B)
1	2
5	6
1	2

$\sigma_{A+B < 5} (R) =$

A	B
1	2
1	2

34

Exemplo: Projeção em Bags

$R($

A,	B)
1	2
5	6
1	2

$\pi_A(R) =$

A
1
5
1

35

Exemplo: Produto em Bags

$R($

A,	B)
1	2
5	6
1	2

$S($

B,	C)
3	4
7	8

$R \times S =$

A	R.B	S.B	C
1	2	3	4
1	2	7	8
5	6	3	4
5	6	7	8
1	2	3	4
1	2	7	8

36

Exemplo: Junção-Theta em Bags

R(A,	B)	S(B,	C)
1		2		3		4	
5		6		7		8	
1		2					

$R \bowtie_{R.B < S.B} S =$

A	R.B	S.B	C
1	2	3	4
1	2	7	8
5	6	7	8
1	2	3	4
1	2	7	8

37

União em Bags

- O número de vezes que um elemento aparecerá na união dos Bags é igual a soma do número de vezes que ele aparece nas duas Bags
- Exemplo:
 - $\{1,2,1\} \cup \{1,1,2,3,1\} = \{1,1,1,1,1,2,2,3\}$

38

Interseção em Bags

- O número de vezes que um elemento aparecerá na intersecção dos Bags é igual ao mínimo entre o número de vezes que ele aparece nas duas Bags
- Example:
 - $\{1,2,1,1\} \cap \{1,2,1,3\} = \{1,1,2\}$.

39

Diferença em Bags

- O nr. de vezes que um elemento aparecerá na em $A - B$ é igual ao nr. de vezes que ele aparece em A menos o nr. de vezes que aparece em B
- Mas nunca é menor que 0!
- Exemplo:
 - $\{1,2,1,1\} - \{1,2,3\} = \{1,1\}$.

40

Cuidados !

- Leis de Bags <> Leis de Conjuntos
- Nem todas as leis algébricas de conjuntos se aplicam!

41

Exemplo

- A união de conjuntos é idempotente
 - $S \cup S = S$.
- Em Bags isso nem sempre se aplica
 - $\{1\} \cup \{1\} = \{1,1\} \neq \{1\}$.

42

Algebra Extendida

δ = elimina duplicados de *bags*.

τ = ordena tuplas.

γ = agrupamento e agregação.

\bowtie = *Junção Externa ou Outerjoin* : evita “dangling tuples” (tuplas pendentes) = tuplas que não são associadas

43

Eliminar Duplicidade

- $R1 := \delta(R2)$.
- R1 consiste de uma cópia de cada tupla que aparece em R2 uma ou mais vezes.

44

Exemplo: Eliminando Duplicidade

$R =$

A	B
1	2
3	4
1	2

$\delta(R) =$

A	B
1	2
3	4

45

Ordenando

- $R1 := \tau_L(R2)$.
 - L é uma lista de alguns atributos de $R2$.
- $R1$ é uma lista de tuplas de $R2$ ordenada primeiramente usando o valor do primeiro atributo de L , então o segundo atributo de L é usado, e assim sucessivamente.
- τ é somente um operador que resulta um conjunto e não uma *bag*.

46

Exemplo: Ordenando

$R =$

A	B
1	2
3	4
5	2

$$\tau_B(R) = [(5,2), (1,2), (3,4)]$$

47

Operação de Agregação

- Operadores de Agregação não são operadores da Álgebra Relacional.
- Na verdade, eles se aplicam para uma coluna inteira de uma tabela, e produzem um resultado único.
- Operadores importantes: SUM, AVG, COUNT, MIN, e MAX.

48

Exemplo: Agregação

$R = (A, B)$

A	B
1	3
3	4
3	2

$SUM(A) = 7$
 $COUNT(A) = 3$
 $MAX(B) = 4$
 $AVG(B) = 3$

49

Operadores para Agrupamento

- $R1 := \gamma_L (R2)$. L é uma lista de elementos que são:
 1. Atributo Individual (*grouping*).
 2. $AGG(A)$, onde AGG é um dos operadores de agregação e A é um atributo.
 - Uma seta e um novo nome de atributo renomeia o componente.

50

Aplicando $\gamma_L(R)$

- Agrupa R de acordo com todos os atributos contidos na lista L .
 - Isto é: forma um grupo para cada lista de valores para estes atributos em R .
- Dentre cada grupo, computa $AGG(A)$ para cada agregação da lista L .
- O resultado tem uma tupla para cada grupo:
 1. Os atributos agrupados e
 2. seu atributo devidamente computado, conforme operador de agregação.

51

Exemplo: Agrupamento/Agregação

$R = (A, B, C)$

A	B	C
1	2	3
4	5	6
1	2	5

$\gamma_{A,B,AVG(C) \rightarrow X}(R) = ??$

Primeiro, agrupa R usando A e B :

A	B	C
1	2	3
1	2	5
4	5	6

AVG
operador de
agregação

Então, média de C
Com os grupos:

A	B	X
1	2	4
4	5	6

52

Exemplo: Junção Externa (Outerjoin)

R =

A	B
1	2
4	5

S =

B	C
2	3
6	7

(1,2) é associada (2,3), mas as outras duas
tuplas ficam pendentes

$R \bowtie S =$

A	B	C
1	2	3
4	5	NULL
NULL	6	7