

Aula 8 – BD1

Álgebra Relacional

Profa. Elaine Faria

UFU - 2018

Introdução

- Linguagens de consulta formais associadas ao modelo relacional
 - Álgebra
 - Usa uma coleção de operadores e cada consulta descreve um procedimento passo a passo para computar o resultado
 - Cálculo
 - Uma consulta descreve a resposta desejada sem especificar como a resposta deve ser computada

Essas linguagens formais tem influenciado linguagens de consulta comerciais, como o SQL

Preliminares

- As entradas e saída de uma consulta são relações
 - Usa-se as instâncias de cada relação de entrada e produz uma instância da relação de saída
- Para referenciar os campos da tabela pode-se usar
 - Nomes de campos → notação mais legível
 - Posição dos campos → mais fácil

Preliminares

Esquema utilizado em nossa aula

Marinheiros(id-marin: integer, nome-marin: string,
avaliação: integer, idade: real)

Barcos(id-barco: integer, nome-barco: string, cor:
string)

Reservas(id-marin: integer, idbarco: integer, dia:
date)

Preliminares

id_marin	nome-marin	avaliação	idade
22	Daniel	7	45
31	Lucas	8	55,5
58	Raul	10	35

Instância M1 de Marinheiros

id_marin	id-barco	dia
22	101	10/10/96
58	103	11/12/96

Instância R1 de Reservas

id_marin	nome-marin	avaliação	idade
28	Ivo	9	35
31	Lucas	8	55,5
44	Gabriel	5	35
58	Raul	10	35

Instância M2 de Marinheiros

Álgebra Relacional

- Consultas
 - Compostas usando uma coleção de operadores
 - Propriedade fundamental do operador
 - Recebe: (uma ou duas) instâncias de relação como argumentos
 - Retorna: uma instância de relação como resultado
- Expressão da álgebra relacional é recursivamente definida como
 - uma relação
 - um operador de álgebra
 - unário aplicado a uma única expressão
 - binário aplicado a duas expressões

Álgebra Relacional

- Operadores
 - Seleção
 - Projeção
 - União
 - Produto cartesiano
 - Diferença

Seleção e Projeção

- Selecionar linhas de uma relação $\rightarrow \sigma$
- Projetar linhas de uma relação $\rightarrow \pi$

Selecionar e projetar permitem a manipulação de dados em uma única relação

Ex: Considere a instância da relação Marinheiros M2

Recuperar as linhas correspondentes aos marinheiros experientes (avaliação maior que 8)

$$\sigma_{\text{avaliação} > 8}(\text{M2})$$

Seleção e Projeção

- Operador de seleção σ
 - Especifica as tuplas a serem mantidas por meio de uma condição de seleção
 - Condição de seleção
 - Combinação Booleana (usa os conectivos \wedge e \vee) de termos que tem a forma
 - $atributo \text{ op } constante$ ou
 - $atributo1 \text{ op } atributo2$
 - Op é um dos operadores de comparação
<, <=, =, \neq , >=, ou >
 - O esquema do resultado de uma seleção é o esquema da instância da relação de entrada

Seleção e Projeção

- Operador de projeção π
 - Permite extrair colunas de uma relação
 - O esquema do resultado de uma projeção é determinado pelos campos que são projetados

Ex: Localizar todos os nomes e avaliações dos marinheiros

$$\pi_{\text{nome-marin,avaliação}}(M2)$$

Ex: Localizar as idades do marinheiros

$$\pi_{\text{idade}}(M2)$$

Embora haja 3 marinheiros com idade 35, uma única tupla idade=35 aparece no resultado da projeção → eliminação de duplicatas

Sistemas reais omitem a tarefa de eliminar duplicatas → processo caro

Seleção e Projeção

- Combinando seleção e projeção

Ex: computar os nomes e avaliações dos marinheiros bem avaliados

$$\pi_{\text{nome-marin,avaliação}}(\sigma_{\text{avaliação} > 8}(M2))$$

Aplica-se a seleção a M2

Aplica-se a projeção a esse resultado

Operações de Conjunto

- Operações sobre conjuntos disponíveis na álgebra
 - União (\cup)
 - Intersecção (\cap)
 - Diferença de conjunto (-)
 - Produto cartesiano (\times)

Operações de Conjunto

- União

- $R \cup M$ retorna uma instância de relação contendo todas as tuplas que ocorrem na instância da relação R , ou na instância da relação M , ou em ambas
- R e M devem ser compatíveis à união
- O esquema do resultado é definido de forma idêntica ao esquema de R

- Duas instâncias são consideradas compatíveis à união se
 - Elas têm o mesmo número de campos e
 - Os campos correspondentes, considerados na ordem da esquerda para a direita têm o mesmo domínio

Operações de Conjunto

- Intersecção

- $R \cap M$ retorna uma instância de relação contendo todas as tuplas que ocorrem em ambas R e M
- As relações R e M devem ser compatíveis à união
- O esquema do resultado é definido de forma idêntica ao esquema de R

- Diferença de conjunto

- $R - M$ retorna uma instância de relação contendo todas as tuplas que ocorrem em R, mas não em M
- As relações R e M devem ser compatíveis à união
- O esquema do resultado é definido de forma idêntica ao esquema de R

Operações de Conjunto

- Produto Cartesiano
 - $R \times M$ retorna uma instância da relação cujo esquema contém os campos de R (na mesma ordem que aparecem em R) seguidos de todos os campos de M (na mesma ordem que eles aparecem em M)
 - O resultado de $R \times M$ contém uma tupla $(r, s) \rightarrow$ concatenação das tuplas r e s para cada par de tuplas $r \in R, s \in M$
 - Os campos de $R \times M$ herdam os nomes dos campos correspondentes de R e M
 - Se R e M contém um ou mais campos com o mesmo nome \rightarrow campos referenciados somente pela posição (conflito de nomeação)

Operações de Conjunto

id_marin	nome-marin	avaliação	idade
22	Daniel	7	45
31	Lucas	8	55,5
58	Raul	10	35
28	Ivo	9	35
44	Gabriel	5	35

União de M1 e M2

id_marin	nome-marin	avaliação	idade
22	Daniel	7	45

Diferença de M1 e M2

id_marin	nome-marin	avaliação	idade
31	Lucas	8	55,5
58	Raul	10	35

Intersecção de M1 e M2

Operações de Conjunto

(id_marin)	nome-marin	avaliação	idade	(id-marin)	id-barco	dia
22	Daniel	7	45	22	101	10/10/96
22	Daniel	7	45	58	103	11/12/96
31	Lucas	8	55,5	22	101	10/10/96
31	Lucas	8	55,5	58	103	11/12/96
58	Raul	10	35	22	101	10/10/96
58	Raul	10	35	58	103	11/12/96

Produto Cartesiano de M1 e R1

Renomear

- Conflitos de nomes podem surgir
 - Ex: M1 x R1
- Operador ρ
 - Usado para renomear explicitamente os campos de uma instância de relação
 - Expressão $\rho(R(\bar{F}), E)$
 - Recebe uma expressão da álgebra relacional E
 - Retorna uma instância de uma relação (nova) R
 - R contém as mesmas tuplas que E e tem o mesmo esquema que E mas alguns campos renomeados
 - Os nomes de campo em R são os mesmos de E exceto os campos renomeados na lista \bar{F}
 - formato nomeantigo \rightarrow nomenovo ou posição \rightarrow nomenovo

Renomear

- Pode-se renomear campos ou relação
 - R e \bar{F} são opcionais

- Exemplo

ρ (C(1 \rightarrow id-marin1, 5 \rightarrow id-marin2), M1 x R1)

Retorna uma relação que tem o esquema

C(id-marin1: integer, nome-marin: string,
avaliação: integer, idade: real, id-marin2:
integer, id-barco: integer, dia: date)

Junções

- Operação usada para combinar informações de duas ou mais relações
 - Pode ser definida como um produto cartesiano seguido de seleções e projeções
 - Aparecem mais comumente que os produtos cartesianos
 - O resultado de um produto cartesiano, em geral, é maior que o de uma junção
- Tipos de junção
 - Junções Condicionais
 - Equijunção
 - Junção Natural

Junções Condicionais

- Argumentos
 - uma condição de junção c e
 - idêntica à condição de seleção
 - um par de instâncias de relação
- Retorna uma instância de relação

$$R \bowtie_c M = \sigma_c (R \times M)$$

c pode referenciar atributos de ambas R e $M \rightarrow R.\text{nome}$ ou $R.i$

$$M1 \bowtie_{M1.\text{id-marin} < R1.\text{id-marin}} R1$$

Equijunção

- Caso especial da operação junção $R \bowtie M$ quando a condição de junção consiste apenas em igualdades (ligadas por \wedge) no formato $R.nome1 = M.nome2$
 - Nesse caso há redundância em manter ambos os atributos no resultado
 - Uma operação de junção é refinada fazendo-se uma projeção adicional na qual $M.nome2$ é excluído
- O esquema do resultado contém os campos de R seguidos pelos campos de M que não aparecem nas condições de junção

$M1 \bowtie_{M1.id-marim=R1.idmarin} R1$

Junção Natural

- Caso especial de junção $R \bowtie M$ na qual as igualdades são especificadas para todos os campos que têm os mesmos nomes em R e em M
 - Pode-se omitir a condição de junção

$$M1 \bowtie_{M1.id-marin=R1.id-marin} R1$$

pode ser denotada por

$$M1 \bowtie R1$$

Se M1 e R1 não tiverem nenhum atributo comum, $M1 \bowtie R1$ é o produto cartesiano

Divisão

- Útil para expressar certos tipos de consulta
 - Ex: “localize os nomes dos marinheiros que reservaram todos os barcos”
- SGBDs não implementam a divisão como um operador distinto
- Ex: Considere duas instâncias de relação A e B, na qual A tem dois campos x e y e B tem o campo y (com o mesmo domínio de A)
 - A/B é o conjunto de todos os valores x tais que, para todo valor y em B, há uma tupla (x,y) em A

Divisão

- Analogia com a divisão de inteiros A e B
 - A/B é o maior inteiro Q tal que $Q \cdot B \leq A$
- Dadas as instâncias de relações A e B
 - A/B é a maior instância da relação Q tal que $Q \times B \subseteq A$
- Ex: seja A uma relação que lista as peças fornecidas pelos fornecedores e B uma relação com as peças listadas
 - A/B_i computa os fornecedores que fornecem todas as peças listadas na instância de relação B_i

Divisão

A

id_f	id_p
F1	P1
F1	P2
F1	P3
F1	P4
F2	P1
F2	P2
F3	P2
F4	P2
F4	P4

B1

id_p
P2

B2

id_p
P2
P4

B3

id_p
P1
P2
p4

A/B1

id_f
F1
F2
F3
F4

A/B2

id_f
F1
F4

A/B3

id_f
F1

id_marin	nome-marin	avaliação	idade
22	Daniel	7	45
29	Bruno	1	33
31	Lucas	8	55,5
32	Alceu	8	25,5
58	Raul	10	35
64	Homero	7	35
71	Ze	10	16
74	Honorato	9	35
85	Americo	3	25,5
95	Bob	3	63,5

Instância M3 de Marinheiros

id_marin	id-barco	dia
22	101	10/10/96
22	102	10/10/98
22	103	10/08/98
22	104	10/07/98
31	102	11/10/98
31	103	11/06/98
31	104	11/12/98
64	101	09/05/98
64	102	09/05/98
74	103	09/08/98

Instância R2 de Reservas

id_barco	nome-barco	cor
101	Interlagos	azul
102	Interlagos	Vermelho
103	Clipper	Verde
104	Marinha	Vermelho

Instância B1 de Barcos

Mais exemplos de consultas

- C1: Encontre os nomes dos marinheiros que reservaram o barco 103

$$\pi_{\text{nome-marin}}((\sigma_{\text{id-barco}=103} \text{Reservas}) \bowtie \text{Marinheiros})$$

Expressando a consulta em partes menores

$$\begin{aligned} & \rho(\text{Temp1}, \sigma_{\text{id-barco}=103} \text{Reservas}) \\ & \rho(\text{Temp2}, \text{Temp1} \bowtie \text{Marinheiros}) \\ & \pi_{\text{nome-marin}}(\text{Temp2}) \end{aligned}$$

Outra solução

$$\pi_{\text{nome-marin}}((\sigma_{\text{id-barco}=103} (\text{Reservas} \bowtie \text{Marinheiros})))$$

Mais exemplos de consultas

- Papel desempenhado pela álgebra em um SGBD relacional
 - Consultas são expressas pelos usuários em uma linguagem tal como a SQL
 - SGBD converte uma consulta SQL em uma forma estendida da álgebra relacional e depois procura outras expressões algébricas que produzem as mesmas respostas, mas que são mais baratas de avaliar

Ex: se a consulta do usuário é escrita como

$$\pi_{\text{nome-marin}}((\sigma_{\text{id-barco}=103}(\text{Reservas} \bowtie \text{Marinheiros})))$$

Um otimizador de consultas encontrará a expressão equivalente

$$\pi_{\text{nome-marin}}((\sigma_{\text{id-barco}=103} \text{Reservas}) \bowtie \text{Marinheiros})$$

Mais exemplos de consultas

- C2: Encontre os nomes dos marinheiros que reservaram um barco vermelho

$\pi_{\text{nome-marin}}((\sigma_{\text{cor}='vermelho'} \text{Barcos}) \bowtie \text{Reservas} \bowtie \text{Marinheiros})$

Expressão equivalente (mais otimizada)

$\pi_{\text{nome-marin}}(\pi_{\text{id-marin}}(\pi_{\text{id-barco}}(\sigma_{\text{cor}='vermelho'} \text{Barcos} \bowtie \text{Reservas}) \bowtie \text{Marinheiros}))$

Mais exemplos de consultas

- C3: Encontre as cores de barcos reservados por Lucas

$\pi_{\text{cor}}((\sigma_{\text{nome-marin}='Lucas'} \text{ Marinheiros}) \bowtie \text{Reservas} \bowtie \text{Barcos})$

- C4: Encontre os nomes dos marinheiros que reservaram pelo menos um barco

$\pi_{\text{nome-marin}}(\text{Marinheiros} \bowtie \text{Reservas})$

Mais exemplos de consultas

- C5: Encontre os nomes dos marinheiros que reservaram um barco vermelho ou um barco verde

$$\rho(\text{Tempbarcos}, (\sigma_{\text{cor}='vermelho'} \text{Barcos}) \cup (\sigma_{\text{cor}='verde'} \text{Barcos}))$$
$$\pi_{\text{nome-marin}}(\text{Tempbarcos} \bowtie \text{Reservas} \bowtie \text{Marinheiros})$$

Ou

$$\rho(\text{Tempbarcos}, (\sigma_{\text{cor}='vermelho' \vee \text{cor}='verde'} \text{Barcos}))$$
$$\pi_{\text{nome-marin}}(\text{Tempbarcos} \bowtie \text{Reservas} \bowtie \text{Marinheiros})$$

Mais exemplos de consultas

- C6: Encontre os nomes dos marinheiros que reservaram um barco vermelho e um barco verde

$\rho(\text{Tempbarcos}, (\sigma_{\text{cor}='vermelho'} \text{Barcos}) \cap (\sigma_{\text{cor}='verde'} \text{Barcos}))$
 $\pi_{\text{nome-marin}}(\text{Tempbarcos} \bowtie \text{Reservas} \bowtie \text{Marinheiros})$

Esta consulta está correta? Por que?

Não. Ela tenta computar marinheiros que reservaram um barco que é vermelho e verde

Mais exemplos de consultas

$\rho(\text{Tempverm}, (\pi_{\text{id-marin}}((\sigma_{\text{cor}='vermelho'} \text{Barcos}) \bowtie \text{Reservas})))$

$\rho(\text{Tempverde}, (\pi_{\text{id-marin}}((\sigma_{\text{cor}='verde'} \text{Barcos}) \bowtie \text{Reservas})))$

$\pi_{\text{nome-marin}}((\text{Tempverm} \cap \text{TempVerde}) \bowtie \text{Marinheiros})$

Reformular esta consulta para encontrar os marinheiros que reservaram barcos vermelhos ou verdes

Mais exemplos de consultas

A C6 pode ser respondida da seguinte forma? Justifique!

ρ (Tempverm, ($\pi_{\text{nome-marin}}((\sigma_{\text{cor}='vermelho'} \text{Barcos}) \bowtie \text{Reservas})$
 $\bowtie \text{Marinheiros}$))

ρ (Tempverde, ($\pi_{\text{nome-marin}}((\sigma_{\text{cor}='verde'} \text{Barcos}) \bowtie \text{Reservas})$
 $\bowtie \text{Marinheiros}$))

$\text{Tempverm} \cap \text{TempVerde}$

Não. Dois marinheiros com o mesmo nome podem reservar barcos vermelho e verde respectivamente. Esse marinheiro será incluído na solução mesmo se ele não tiver reservado um barco verde e vermelho

Mais exemplos de consultas

- C7: Encontre o nome dos marinheiros que reservaram pelo menos dois barcos

ρ (Tempreservas $(\pi_{\text{id-marin}, \text{nome-marin}, \text{id-barco}}(\text{Marinheiros} \bowtie \text{Reservas}))$

ρ (Paresreservas, $(1 \rightarrow \text{id-marin1}, 2 \rightarrow \text{nome-marin1}, 3 \rightarrow \text{id-barco1}, 4 \rightarrow \text{id-marin2}, 5 \rightarrow \text{nome-marin2}, 6 \rightarrow \text{id-barco2}),$
Tempreservas X Tempreservas)

$\pi_{\text{nome-marin1}} \sigma_{(\text{id-marin1}=\text{id-marin2}) \wedge (\text{id-barco1} \neq \text{id-barco2})} \text{Paresreservas}$

Mais exemplos de consultas

- C8: Encontre os id-marins dos marinheiros com idade acima de 20 que não reservaram um barco vermelho

$\pi_{\text{id-marin}}(\sigma_{\text{idade} > 20} \text{Marinheiros}) -$
 $\pi_{\text{id-marin}}((\sigma_{\text{cor}='vermelho'} \text{Barcos}) \bowtie \text{Reservas} \bowtie \text{Marinheiros})$

Mais exemplos de consultas

- C9: Encontre os nomes dos marinheiros que reservaram todos os barcos

ρ (Tempidmarin, $(\pi_{\text{id-marin, id-barco}} \text{Reservas}) / (\pi_{\text{id-barco}} \text{Barcos})$)
 $\pi_{\text{nome-marin}}(\text{Tempidmarin} \bowtie \text{Marinheiros})$

Mais exemplos de consultas

- C10: Encontre os nomes dos marinheiros que reservaram todos os barcos de nome Interlagos

$$\rho(\text{Tempidmarin}, (\pi_{\text{id-marin, id-barco}} \text{Reservas}) /$$
$$(\pi_{\text{id-barco}} (\sigma_{\text{nome-barco}='Interlagos'} \text{Barcos})))$$
$$\pi_{\text{nome-marin}}(\text{Tempidmarin} \bowtie \text{Marinheiros})$$

Referências

- R. Ramakrishnan e J. Gehrke, *Database Management Systems*, 3a Edição, McGraw-Hill, 2003.