

Desda página 89 até página 108 do livro *Database System Concepts 4th edition*

Estrutura básica

- Dados os conjuntos $D_1, D_2, \dots, D_n$ uma **relação** r é um **subconjunto** de $(D_1 \times D_2 \times \dots \times D_n)$
- Ou seja, a **relação** é um **conjunto de tuplos** $(a_1, a_2, \dots, a_n)$ em que $a_i \in D_i$

Linguagem de Consulta

- Linguagem a que o utilizador recorre para obter informação a partir da base de dados
- Existem Categorias de Linguagens (implementadas)
 - Procedimentais
 - Declarativas
- Existem Linguagens Puras (conceptuais)
 - Álgebra Relacional
- As linguagens puras formam a base teórica das linguagens de consulta utilizadas na prática
- Uma pergunta é uma “função” /operação que, dado um conjunto de relações, devolve uma relação
- Portanto, uma resposta da base de dados é uma relação
- Mas que operadores podemos usar?

Álgebra Relacional

- Seis operadores básicos
 - Selecção
 - $\sigma_p(r)$
 - p é o predicado a verificar pela operação de selecção
 - $=, \neq, >, \geq, <, \leq$
 - Devolve uma relação contendo os tuplos da relação r que verifiquem o predicado p
 - É uma operação comotativa
 - Projecção
 - $\pi_{A_1, A_2, \dots, A_k}(r)$
 - r é uma relação
 - $A_1..A_k$ são os atributos projectados (mostrados)

- O resultado é um conjunto de todos os tuplos da relação R, apenas exibindo os atributos **A1...Ak**
- É comotativa com a selecção
- União
 - $r \cup s$
 - r, s são relações que devem ter a mesma aridade
 - O esquema das relações devem ser compatíveis
 - Devolve uma relação com os tuplos de r e os de s
- Diferença de Conjuntos
 - $r - s$
 - r, s são relações que devem ter a mesma aridade
 - Os domínios dos atributos de r e s devem ser compatíveis
 - Devolve uma relação contendo os tuplos de r excepto os de s
- Produto Cartesiano
 - $r \times s$
 - Assume que os atributos de r(R) e s(S) são disjuntos
 - Se os atributos de r(R) e s(S) não forem disjuntos, então têm que ser utilizadas renomeações
 - Devolve uma relação com todos os tuplos compostos por cada tudo de cada relação
- Renomeação
 - $\rho_{x(A_1, A_2, \dots, A_n)}(E)$
 - Devolve uma expressão chamada X com o mesmo domínio do esquema de relação, com os atributos renomeados (A1, A2, ..., An)
 - Se não tiver nomeação de atributos, os atributos da relação resultante terão o mesmo nome que da relação de entrada
- Têm como argumento relações de entrada e devolvem uma relação com resultado

Operações estendidas da álgebra relacional

- Definem-se outras operações que não aumentam a expressividade da álgebra relacional, mas que simplificam algumas consultas habituais
 - Intersecção de conjuntos
 - $r \cap s$
 - As relações r e s devem ter a mesma aridade
 - Os atributos devem ser compatíveis
 - Devolve uma relação contendo os tuplos comuns aos dois
 - $= r - (r - s) = s - (s - r)$
 - Junção Natural
 - $r \bowtie s$
 - Sejam r e s relações. A relação resultante tem o esquema de relação $r(R) \cup s(S)$

- Sejam t_r e t_s tuplos com o mesmo valor em cada atributo de $R \cap S$
- Se $R \cap S = \{ \}$ então $r(R) \bowtie s(S) = r(R) \times s(S)$
- O tuplo t faz parte da relação resultante
- É comotativa e associativa
- Divisão
 - $r \div s$
 - Adequa-se para consultas que incluam a frase “para todo”
 - Sejam r e s relações em que o $s(S)$ é um subconjunto/subesquema de $r(R)$
 - O resultado é uma relação no esquema $R-S$
 - Cada tuplo de R que está associado a todo o tuplo de S faz parte da relação resultante, num esquema $R-S$
 - $= \pi_{R-S}(r) - \pi_{R-S}((\pi_{R-S}(r) \times s) - \pi_{R-S,s}(r))$
 - Porquê?
 - $\pi_{R-S}(r) \times s$ dá os elementos de r com todos os valores de S
 - $\pi_{R-S,s}(r)$ reordena os atributos de r
 - $\pi_{R-S}(\pi_{R-S}(r) \times s) - \pi_{R-S,s}(r)$ dá os tuplos t em $\pi_{R-S}(r)$ tal que para algum tuplo $u \in s$, $tu \notin r$.
 - Atribuição
 - $temp1 \leftarrow op$
 - Atribui um identificador a uma determinada operação
 - Pode ser utilizado nas expressões seguintes
 - Aumentam a expressividade da Álgebra Relacional
 - Projecção Generalizada
 - $\pi F_1, F_2 \dots F_n (E)$
 - Permite a utilização de funções aritméticas na lista de projecção
 - E é uma expressão de álgebra relacional
 - Cada expressão $F_1 \dots F_n$ é uma expressão aritmética envolvendo constantes e atributos do esquema de E
 - Funções de Agregação
 - Aplicam-se a uma colecção de valores e devolvem um único valor como resultado
 - avg, min, max, sum count
 - **avg ignora os valores nulos!**
 - **sum ignora valores nulos!**
 - **count não ignora valores nulos!**
 - $G_1, G_2, \dots, G_n \mathop{\mathcal{G}} F_1(A_1), F_2(A_2), \dots, F_n(A_n)(E)$
 - E é uma expressão de álgebra relacional

- $G_1, \dots, G_n$ é uma lista de atributos de agrupamento (pode ser vazia)
 - Cada F_i é uma função de agregação
 - Cada A_i é um nome de um atributo
 - Agrupa por G 's e aplica as funções F 's
- Junção Externa
 - Uma extensão da operação de junção que evita a perda de informação
 - Calcula a junção e depois adiciona ao resultado os tuplos de uma relação que não estão relacionados com a outra relação na junção
 - Utiliza valores nulos
 - *null* significa que o valor é desconhecido ou que não existe
 - Simplificadamente, todas as comparações com null são falsas por definição

contas			titulares	
número-conta	bação	saldo	nome-cliente	número-conta
A-102	Sete Rios	400	Joana Sobral	A-102
A-202	Benfica	900	Pedro Silva	A-202
A-216	Almada	750	Ana Dias	A-310

■ Junção externa esquerda

contas ⋈ titulares			
número-conta	bação	saldo	nome-cliente
A-102	Sete Rios	400	Joana Sobral
A-202	Benfica	900	Pedro Silva
A-216	Almada	750	NULL

contas			titulares	
número-conta	bação	saldo	nome-cliente	número-conta
A-102	Sete Rios	400	Joana Sobral	A-102
A-202	Benfica	900	Pedro Silva	A-202
A-216	Almada	750	Ana Dias	A-310

■ Junção externa direita

Contas ⋈ titulares			
número-conta	bação	saldo	nome-cliente
A-102	Sete Rios	400	Joana Sobral
A-202	Benfica	900	Pedro Silva
A-310	NULL	NULL	Ana Dias

contas			titulares	
número-conta	bação	saldo	nome-cliente	número-conta
A-102	Sete Rios	400	Joana Sobral	A-102
A-202	Benfica	900	Pedro Silva	A-202
A-216	Almada	750	Ana Dias	A-310

■ Junção externa total

contas			titulares
número-conta	bação	saldo	nome-cliente
A-102	Sete Rios	400	Joana Sobral
A-202	Benfica	900	Pedro Silva
A-216	Almada	750	NULL
A-310	NULL	NULL	Ana Dias

Valores nulos

- Comparações com valores nulos devolvem o valor de verdade **unknown**
- Lógica a três valores com o valor de verdade **unknown**
 - Or
 - Unknown OR true = true
 - Unknown OR false = unknown
 - Unknown OR unknown = unknown
 - And
 - Unknown AND true = unknown
 - Unknown AND false = false
 - Unknown AND unknown = unknown
 - Not
 - Not unknown = unknown
 - Em SQL “P is unknown” é verdade sse P tem o valor de verdade unknown
- Predicado da selecção é tratado como false se tiver valor de verdade unknown

Modificação de Base de dados

- O conteúdo da base de dados pode ser modificado através das seguintes operações
 - Remoção
 - Expressa de maneira semelhante a uma consulta
 - Porém, os tuplos seleccionados são removidos

- $r \leftarrow r - E$, onde r é uma relação e E é uma operação de álgebra relacional

○ Inserção

- Para inserir informação podemos:
 - Especificar um tuplo a ser inserido
 - Escrever uma consulta cujo resultado é um conjunto de tuplos a inserir
- Na álgebra relacional, uma inserção é expressa por
 - $r \leftarrow r \cup E$, onde r é uma relação e E é uma expressão

- Inserir informação na base de dados especificando que um novo paciente, com BI nº 10000000 e nome Paulo, teve uma consulta (nº1000) no dia 30-09-2003 com o médico João (assumindo que só há um médico com esse nome)

$pacientes \leftarrow pacientes \cup \{(10000000, 'Paulo', null, null, null)\}$
 $consultas \leftarrow consultas \cup$
 $\cup \Pi_{1000, 30-09-2003, 1000, nEmpr} (\sigma_{nomeM = "João"}(médicos))$

○ Actualização

- Mecanismo para alterar um valor de um tuplo sem alterar todos os valores do tuplo
- Recorre-se ao operador de projecção generalizada para efectuar este tipo de tarefa
 - $r \leftarrow \prod F_1, F_2 \dots F_n (r)$
- Cada F_i ou F_i é o i -ésimo atributo de r , se o i -ésimo atributo não for alterado, ou
- Uma expressão F_i , envolvendo apenas constantes e atributos de r , que permite calcular o novo valor do atributo

- Pague juros de 5% em todas as contas

$account \leftarrow \prod_{AN, BN, BAL} * 1.05 (account)$

em que AN , BN e BAL são abreviaturas para *account-number*, *branch-name* e *balance*, respectivamente.

- Pague 6% de juros em todas as contas com saldo superior a \$10,000 e 5% às restantes contas.

$account \leftarrow \prod_{AN, BN, BAL} * 1.06 (\sigma_{BAL > 10000} (account))$
 $\cup \prod_{AN, BN, BAL} * 1.05 (\sigma_{BAL \leq 10000} (account))$