

# Lenguajes de Consulta

Fuentes <a href="http://www-db.stanford.edu/~ullman/fcdb.html">http://www-db.stanford.edu/~ullman/fcdb.html</a>
Fundamentals of Database Systems, Elmasri y Navathe
Database System Concepts, Silberschatz, Korth, Sudarshan



# Lenguajes

- Permiten recuperación de datos de la bd
- Varios lenguajes para MR
  - Lenguajes formales
  - Basados en lógica y algebra
  - Facilita optimización
- Lenguajes
  - Acceso eficiente a bd
  - No cálculos complejos
  - Declarativos vs. imperativos



### Lenguajes de Consulta Relacional

- □ Base implementación SQL
- Algebra Relacional
  - Imperativa
  - Útil representación planes de consulta
- □ Cálculo Relacional
  - Declarativa
  - Enfasis QUÉ y no en CÓMO





### Lenguajes de Consulta Relacional

- Consultas aplicables a instancias de relación
- Producen instancias relación
- Esquema relaciones entrada fijo
- Esquemas basados en posición vs.
   basados en nombres de atributos



### **ALGEBRA RELACIONAL**





# Algebra Relacional

#### □ Algebra

- Operandos: relaciones o variables de relación
- Operadores: aplicados sobre relaciones
- Operaciones se pueden componer

# Algebra usada como lenguaje de consulta para las relaciones





# Algebra Relacional: Core

- Selección (σ) Selecciona un subconjunto de filas o tuplas de una relación (horizontal)
- Proyección (π) Conserva algunas columnas de la relación (vertical)
- Producto-cartesiano (x) Permite combinar dos relaciones
- □ *Diferencia-conjuntos* ( ) tuplas en r1 no en r2.
- □ *Union* ( ∪ ) Tuplas en r1 o en r2
- $\square$  Renombramiento  $(\rho)$



# Algebra Relacional

#### Operadores Unarios

- Selección (σ)
- Proyección (π)
- Renombramiento (ρ)

#### Operaciones de conjunto

- Producto-cartesiano (x)
- Diferencia-conjuntos ( )
- **■** *Union* ( ∪ )
- Intersección ( ∩ )



# Algebra Relacional

#### Operadores Relacionales Binarios

- *Join* (⋈)
- División

#### Otras Operaciones

- OUTER JOIN, OUTER UNION
- Funciones Agregadas (SUM, COUNT, AVG, MIN, MAX)



$$R1 := \sigma_C(R2)$$

- C es una condición( predicado) aplicable a attributes de R2
- R1 contiene todas las tuplas de R2 que satisfacen C
- Esquemas R1 y R2 iguales



$$\sigma_c(r) = \{t \mid t \in r \ y \ c(t)\}\$$

c es una fórmula del cálculo proposicional formada por términos y operadores  $\land$  (and),  $\lor$  (or),  $\neg$  (not)

Cada término puede ser:

<atributo> op <atributo> o <constante>

donde *op* puede ser: =,  $\neq$ , >,  $\geq$ , <,  $\leq$ 



### Selección

#### Relación r

Α	В	С	D
α	α	1	7
α	β	5	7
β	β	12	3
β	β	23	10

$$\sigma_{A=B \wedge D > 5}$$
 (r)

Α	В	С	D
α	α	1	7
β	$\beta$	23	10



# Ejemplo

#### Relación Matrícula

estudiante	curso	calif
Luis	BD	3.5
José	MDI	4.7
Maria	BD	3.5
Luis	MDI	4.0

Tabulado-Luis= **o**estudiante="Luis" (Matricula)

estudiante	curso	calif
Luis	BD	3.5
Luis	MDI	4.0



Por conmutatividad una secuencia (cascada) de operaciones SELECT se puede aplicar en cualquier orden

$$\sigma_{\text{cond1}>}(\sigma_{\text{cond2}>}(\sigma_{\text{cond3}>}(R)) = \\ \sigma_{\text{cond2}>}(\sigma_{\text{cond3}>}(\sigma_{\text{cond1}>}(R)))$$



Una secuencia de operaciones SELECT se puede reemplazar por la conjunción de condiciones

$$\sigma_{\text{cond1>}}(\sigma_{\text{cond2>}}(\sigma_{\text{cond3>}}(R)) =$$

$$\sigma_{\text{cond1>AND< cond2>AND < cond3>}}(R)))$$



El número de tuplas en el resultado de un SELECT es menor que (o igual a) el número de tuplas de la relación de entrada R





# Proyección π

$$R1 := \pi_L(R2)$$

- L lista de atributos (A<sub>1</sub>,A<sub>2</sub>,...,A<sub>n</sub>) del esquema
   R2
- R1 está formada por tuplas de R2, con los atributos en el orden de la lista L
- tuplas duplicadas se eliminan si hay alguna (conjuntos no tienen elementos duplicados)

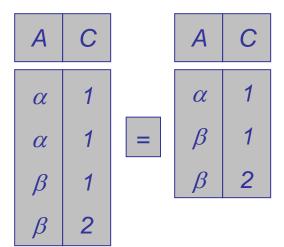


# Proyección

#### Relación r

Α	В	С
α	10	1
α	20	1
β	30	1
β	40	2

$$\prod_{A,C} (r)$$





# Proyección: Propiedades

El total de tuplas resultado de  $\pi_{<list>}(R)$  es siempre menor o igual al total de tuplas de R

- Si la lista de atributos incluye una llave de R, el número de tuplas en el resultado de la proyección es igual al total de tuplas en R
- PROYECCIÓN no es commutativa

$$\pi_{<\text{list}1>}(\pi_{<\text{list}2>}(R)) = \pi_{<\text{list}1>}(R)$$
 sobre conteniendo los atributos en t1>



# Ejemplo

#### Relación Matricula

estudiante	curso	calif
Luis	BD	3.5
José	MDI	4.7
Maria	BD	3.5
Luis	MDI	4.0

Estudiantes =  $\pi_{\text{estudiante}}$  (Matricula)

estudiante Luis José María





# Proyección Extendida

La lista L, en  $\pi_L$ , puede contener expresiones aplicables a los atributos:

- Aritméticas como A+B → C
- Ocurrencias duplicadas sobre el mismo atributo



# Ejemplo

$$\Pi_{A+B\to C,A}(R) = \begin{array}{c|c} C & A \\ \hline 3 & 1 \\ \hline 7 & 3 \end{array}$$



#### Producto Cartesiano

#### $R3 := R1 \times R2$

- Cada tupla t1 de R1 se combina con cada tupla t2 de R2
- La concatenación t1t2 es una tupla de R3
- El esquema de R3 se forma con los atributos de R1 y de R2 en el orden en que aparecen
- Si el mismo atributo A aparece en ambas relaciones R1 y R2, usar R1.A y R2.A





# Ejemplo $R3 = R1 \times R2$

R1

Α	В
1	2
3	4

R2

В	C
5	6
7	8
9	10

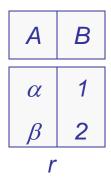
**R3** 

Α	R1.B	R2.B	С
1	2	5	6
1	2	7	8
1	2	9	10
3	4	5	6
3	4	7	8
3	4	9	10



### Producto Cartesiano

#### Relaciones



С	D	E
α	10	а
$\beta$	10	a
$\beta$	20	b
γ	10	b
		S

rxs

A	В	С	D	Ε
$\alpha$	1	α	10	а
$\alpha$	1	$\beta$	10	a
$\alpha$	1	$\beta$	20	b
$\alpha$	1	γ	10	b
$\beta$	2	$\alpha$	10	a
$\beta$	2	$\beta$	10	а
$\beta$	2	$\beta$	20	b
$\beta$	2	$\gamma$	10	b



#### Producto Cartesiano

$$r \times s = \{t \mid q \mid t \in r \land q \in s\}$$

- Se asume que los atributos en r(R) y
   s(S) son disjuntos (R ∩ S = Ø)
- Si (R ∩ S ≠ Ø) necesidad de usar renombramiento



# Renombramiento p

- Operador ρ produce un nuevo esquema de relación
- □ R1 :=  $\rho_{R1(A1,...,An)}(R2)$  construye relación R1 con atributos  $A_1,...,A_n$  y las mismas tuplas de R2
- Notación simplificada

$$R1(A1,...,An) := R2$$



### Renombramiento

- Permite nombrar atributos o relaciones
- □ Formas de renombramiento
  - ρ<sub>S (B1, B2, ..., Bn )</sub>(R) cambia
    - Por S el nombre de la relación
    - por B1, B1, .....Bn los nombres de los aributos o columnas





#### Renombramiento

#### □ Formas de renombramiento

- ρ<sub>S</sub>(R) cambia
  - Cambia únicamente por S el nombre de la relación
- ρ<sub>(B1, B2, ..., Bn)</sub>(R)
  - Cambia unicamente los nombres de las columnas o atributos por B1, B1, .....Bn



### Unión U

#### $\square$ Operador binario $r \cup s$

$$r \cup s = \{t \mid t \in r \lor t \in s\}$$

- r, s deben tener la misma aridad
- Devuelve tuplas que están en r, s ó ambas
- Tuplas duplicadas se eliminan
- Dominios de atributos compatibles (mismo tipo de valores columnas correspondientes)
- Relaciones UNION compatibles



# Relaciones Unión Compatibles

Dos relaciones r y r'con esquemas R(A<sub>i</sub>:D<sub>i</sub>), R'(A<sub>i</sub>:D<sub>i</sub>) y cardinalidades m y m', respectivamente, son compatibles cuando ambas están definidas sobre el mismo dominio.

$$\forall A_i \exists A'_j | dom(A_i) = dom(A'_j)$$
  
 $\forall A'_j \exists A_j | dom(A'_j) = dom(A_j)$ 



# Unión U

#### Relaciones

Α	В
α	1
α	2
β	1

Α	В	
α	2	
β	3	
S		

rus

A	В
α	1
α	2
β	1
β	3



### Observación

Cuando los esquemas R y R'no se corresponden (el nombre de los atributos son diferentes en las relaciones, o están en distinto orden) se sugiere renombrar los atributos en la relación resultante





# Ejemplo

# Calcular la unión de las siguientes relaciones

#### M

Cedula-M	Nombre-M
3232330	María Ruíz
3455775	Rosa Toro
3216789	Julia Muñoz

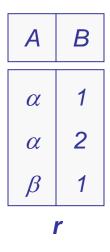
Η

Nombre-H	Cedula-H	
6756989	Luis Lara	
6453446	Jose Cruz	



### Diferencia -

#### Relaciones





#### Diferencia

$$r-s = \{t \mid t \in r \land t \notin s\}$$

- Relaciones UNION compatibles
  - ry s deben tener la misma aridad
  - Dominios de atributos de r y s deben ser compatibles



# Composición de Operaciones

Expresiones con múltiples operadores

$$\Box \sigma_{A=C}(r \times s)$$

$$\sigma_{A=C}(r \times s)$$

A	В	С	D	Ε
α	1	$\alpha$	10	а
$\boldsymbol{\beta}$	2	$\beta$	10	a
$\boldsymbol{\beta}$	2	$\beta$	20	b

Α	В	С	D	E
α	1	α	10	а
$\alpha$	1	$\beta$	10	a
$\alpha$	1	$\beta$	20	b
$\alpha$	1	γ	10	b
$\beta$	2	$\alpha$	10	a
$\beta$	2	$\beta$	10	a
β	2	$\beta$	20	b
$\beta$	2	γ	10	b



### Definición Formal

Sean  $E_1$  y  $E_2$  dos expresiones relacionales Las siguientes son expresiones del álgebra

- *E*<sub>1</sub> ∪ *E*<sub>2</sub>
- $E_1 E_2$
- $E_1 \times E_2$
- $\sigma_p(E_1)$ , p es un predicado sobre los atributos en  $E_1$
- $\Pi_{\mathcal{S}}(E_1)$ , S lista de atributos en  $E_1$
- $\rho_X(E_1)$ , x es un nombre nuevo para el resultado de  $E_1$



