

Tema 9: Introducción al álgebra relacional

Introducción a la Ingeniería del Software y los Sistemas de Información I
Ingeniería Informática – Tecnologías Informáticas
Departamento de Lenguajes y Sistemas Informáticos



1. ¿Qué es el álgebra relacional?
2. Conceptos previos
3. Operadores conjuntistas
4. Operadores relacionales

¿Qué es el Álgebra Relacional (AR)?

- Es un conjunto de **operadores** sobre relaciones propuesto por Codd que permiten expresar **consultas** sobre un modelo relacional.
- Los operadores del AR devuelven como resultado una **relación derivada**, por lo que pueden **anidarse** formando expresiones complejas.

Operadores conjuntistas

- Unión $R \cup S$
- Intersección $R \cap S$
- Diferencia $R - S$
- Producto cartesiano $R \times S$

Operadores relacionales

- Selección $\sigma_f(R)$
- Proyección $\Pi_B(R)$
- Combinación $R \bowtie S$
- División $R \div S$
- Agregación $\gamma_G^F(R)$

- Nombres de atributos: se prefijan con el nombre de la relación a la que pertenecen cuando pueda haber **ambigüedad** (Notación calificada)

$R.a$ = atributo a de la relación R

- Relaciones compatibles
 - Son relaciones cuyas **intensiones comparten los mismos dominios**, de modo que cada atributo tiene su correspondiente en la otra relación y ambos están definidos sobre el mismo dominio.
 - Sobre ellas se pueden aplicar los operadores conjuntistas de **unión, diferencia e intersección**.
 - Si se quiere aplicar dichos operadores sobre relaciones no compatibles por tener nombres de atributos diferentes, es necesario usar el operador de **renombrado** para hacerlas compatibles.

$T = \rho_{a/b}(R)$: el atributo $R.b$ pasa a nombrarse como $T.a$

Unión: $T \leftarrow R \cup S$

- $I(T) = I(R)$
- $E(T) = E(R) \cup E(S)$

Intersección: $T \leftarrow R \cap S$

- $I(T) = I(R)$
- $E(T) = E(R) \cap E(S)$

Diferencia: $T \leftarrow R - S$

- $I(T) = I(R)$
- $E(T) = E(R) - E(S)$

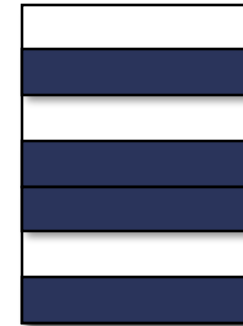
Producto cartesiano: $T \leftarrow R \times S$

- $I(T) = I(R) \cup I(S)$
- $E(T) = \{(r, s) | r \in E(R) \wedge s \in E(S)\}$

Selección: $T \leftarrow \sigma_f(R)$

- $I(T) = I(R)$
- $E(T) = \{r \in E(R) | f(r)\}$
- f es una fórmula bien formada sobre $I(R)$

Selecciona las tuplas de R que satisfacen f

A diagram of a table with 6 rows and 1 column. The rows alternate between white and dark blue, starting and ending with a white row. This represents the selection operation where only rows satisfying a condition are kept.

! PK

AK

@ FK

Usuarios (!usuariold, #nombre, dirección, teléfono, población)

Productos (!productold, descripción, precio)

Pedidos (!pedidold, @ usuariold, @ productold, cantidad)

Usuarios				
usuariold	nombre	dirección	téléfono	provincia
U1	David	Dirección David	Teléfono David	Sevilla
U2	Marta	Dirección Marta	Teléfono Marta	Málaga
U3	Pedro	Dirección Pedro	Teléfono Pedro	Barcelona
...				

Productos		
productold	descripción	precio
P1	Xiaomi mi band 4	36€
P2	Motorola One	399€
P3	Correa compatible mi band 4	10€
...		

Pedidos			
pedidold	usuariold	productold	Cantidad
P1	U1	P1	2
P2	U1	P3	2
P3	U2	P2	1
...			

1. Usuarios de Sevilla

$$Sevillanos \leftarrow \sigma_{provincia="Sevilla"}(Usuarios)$$

2. Usuarios que no son de Sevilla

$$Resultado \leftarrow \sigma_{provincia \neq "Sevilla"}(Usuarios)$$

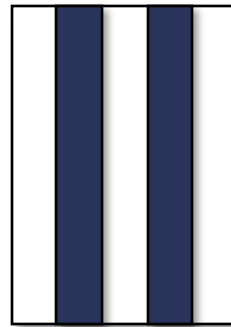
3. Malagueños y sevillanos

$$Malagueños \leftarrow \sigma_{provincia="Málaga"}(Usuarios)$$
$$Resultado \leftarrow Malagueños \cup Sevillanos$$
$$Resultado' \leftarrow \sigma_{provincia="Málaga" \text{ OR } provincia="Sevilla"}(Usuarios)$$

Proyección: $T \leftarrow \Pi_B(R)$, siendo $B \subseteq I(R)$

- $I(T) = B$
- $E(T) = \{t | \forall b \in B \wedge \exists r \in E(R) \wedge t.b = r.b\}$

Extrae las columnas B de R



1. Nombre de los usuarios de Málaga y Sevilla

$Malagueños \leftarrow \sigma_{provincia="Malaga"}(Usuarios)$

$Sevillanos \leftarrow \sigma_{provincia="Sevilla"}(Usuarios)$

$SM \leftarrow Malagueños \cup Sevillanos$

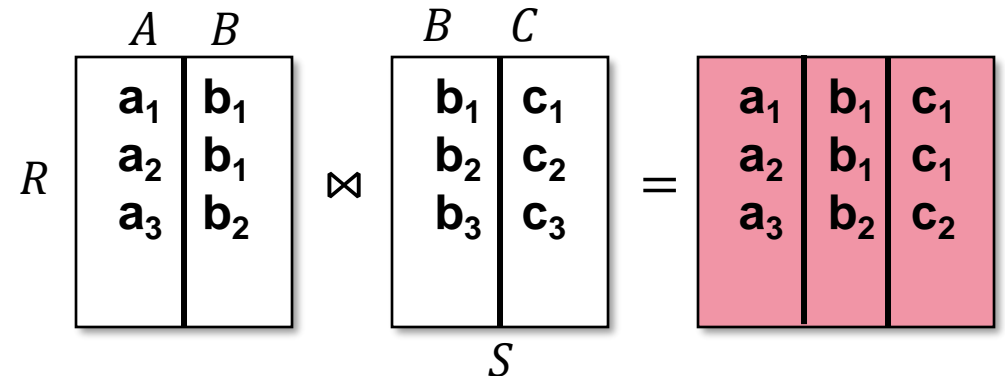
$Resultado \leftarrow \Pi_{nombre}(SM)$

Combinación natural (Natural Join): $T \leftarrow R \bowtie S$, siendo $I(R) \cap I(S) = C \neq \emptyset$

- $I(T) = I(R) \cup I(S)$
- $E(T) = \{r \cup s \mid r \in E(R) \wedge s \in E(S) \wedge \forall c \in C \models r.c = s.c\}$
- Combina la proyección, la selección y el producto cartesiano en una sola operación.
- Crea una relación que **no** está normalizada.
- El resultado es el producto cartesiano $R \times S$ seleccionando las tuplas en las que las claves primarias de R son iguales a las claves ajenas de S

Natural Join

- Permite combinar los datos de relaciones enlazadas mediante **claves ajenas**.
- Es un operador derivado:
 - $\{c_i\} = I(R) \cap I(S)$ (atributos comunes)
 - $\{u_j\} = I(R) \cup I(S)$ (atributos sin repetidos)
 - $R \bowtie S = \Pi_{u_j}(\sigma_{R.c_i=S.c_i}(R \times S))$



Nombre de los usuarios, descripción del producto y cantidad solicitada

$$UPP \leftarrow Usuarios \bowtie Pedidos \bowtie Productos$$

$$Resultado \leftarrow \Pi_{\begin{matrix} Usuarios.nombre, \\ Productos.descripcion, \\ Pedidos.cantidad \end{matrix}} (UPP)$$

Nombre de los usuarios que no han realizado ningún pedido

$$Nombres \leftarrow \Pi_{nombre}(Usuarios)$$

$$NombresConPedido \leftarrow \Pi_{nombre}(Usuarios \bowtie Pedidos)$$

$$Resultado \leftarrow Nombres - NombresConPedido$$

División: $T \leftarrow R \div S$, siendo $I(S) \subset I(R)$

- $I(T) = I(R) - I(S)$
 - $E(T) = \{t | \forall s \in E(S) \models ts \in E(R)\}$
- Sean $R(r,s)$ y $S(s)$ dos relaciones, la división R/S devuelve los valores de r tales que para todo valor de s en S existe una tupla (r,s) en R .

a	x
a	y
a	z
b	x
c	y

÷

x
y

=

a

- El operador de división es un operador derivado.
- Sean $R(r,s)$ y $S(s)$, entonces $R \div S$:

$allComb \leftarrow (\Pi_r(R)) \times S$ // Todas las combinaciones posibles de r y s

$notAllComb \leftarrow allComb - R$ // Combinaciones que faltan

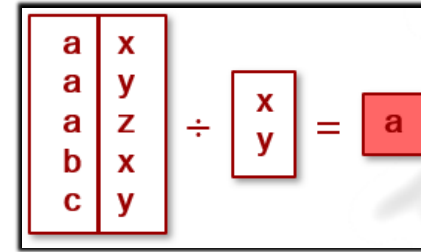
$onlyR \leftarrow \Pi_r(notAllComb)$ // Los r que no sirven

$Resultado \leftarrow \Pi_r(R) - onlyR$ // Los r que sí tienen todas las combinaciones

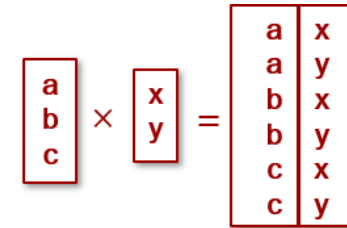
$$R \div S \leftarrow \Pi_r(R) - (\Pi_r((\Pi_r(R)) \times S) - R)$$

Operador de división

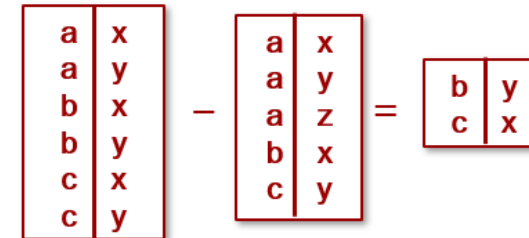
$R(r,s)$ y $S(s)$, $R \div S$



$allComb \leftarrow (\Pi_r(R)) \times S$ //



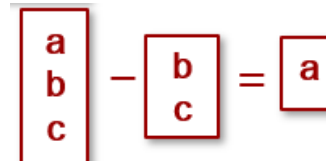
$notAllComb \leftarrow allComb - R$ //



$onlyR \leftarrow \Pi_r(notAllComb)$ //



Resultado $\leftarrow \Pi_r(R) - onlyR$ //



Nombre de los usuarios que han comprado **todos** los productos

$$ProdsId \leftarrow \Pi_{productId}(Productos)$$
$$UPP \leftarrow Usuarios \bowtie Pedidos \bowtie Productos$$
$$NDC \leftarrow \Pi_{\substack{Usuarios.nombre \\ Productos.productId}}(UPP)$$
$$Resultado \leftarrow NDC \div ProdsId$$

Agregación: $\gamma_G^F(R)$, siendo:

- $G \subset I(R)$
 - Determina el número de particiones por cada valor $g \in G$. Si $G = \emptyset$ la partición es toda la relación.
- F , función de agregación definida sobre R
 - F no puede ser vacío.
 - Las funciones de agregación (*count*, *sum*, *min*, *max*, *avg*, ...) tienen un sólo parámetro.
 - F puede contener también atributos que estén en G .*

* En esos casos, la función de agregación aplicada sería la función identidad.

Agregación: $\gamma_G^F(R)$:

$\gamma^{count(*)}(R)$
count(*)
5

$\gamma^{max(X)}(R)$
max(X)
15

Una partición $G=\emptyset$
Una fila resultado

R		
ID	A	X
1	A	10
2	b	5
3	a	8
4	b	9
5	c	15

$\gamma_A^?(R)$	
A	X
a	10
a	8
b	5
b	9
c	15

Tres particiones

Tres tuplas
en el resultado

$\gamma_A^{A,sum(X)}(R)$	
A	sum(X)
a	18
b	14
c	15

- Obtener el nombre de usuario junto con el importe total de productos pedidos

$$UPP \leftarrow Usuarios \bowtie Pedidos \bowtie Productos$$
$$Resultado \leftarrow \gamma_{nombre}^{nombre, sum(cantidad*precio)}(UPP)$$

- Obtener la descripción del producto junto con el número de pedidos

$$Resultado \leftarrow \gamma_{descripcion}^{descripcion, count(idPedido)}(UPP)$$

Tema 9: Introducción al álgebra relacional

Introducción a la Ingeniería del Software y los Sistemas de Información I
Ingeniería Informática – Tecnologías Informáticas
Departamento de Lenguajes y Sistemas Informáticos

