# TEMA 6

# FORMALIZACIÓN DEL ÁLGEBRA RELACIONAL

## Nomenclatura:

Dado un esquema de relación R(A) definido como:

$R\ (A) = R\ (A_1 : D_1, A_2 : D_2, ..., A_n : D_n)$

la relación r(R), definida sobre el esquema R, de grado n y cardinalidad m estará constituida por un conjunto de m tuplas:

$r(R) = \{ t_i \}^{m}_{i=1}$ donde $t_i = < v_{i1}, v_{i2}, ...., v_{in} > \mid v_{ij} \in D_j$

## Compatibilidad de esquemas:

Se dice que dos esquemas de relación $R1\ (A1_i : D1_i)$ y $R2\ (A2_i : D2_i)$, con cardinalidades $m_1$ y $m_2$, son compatibles cuando ambos están definidos sobre el mismo conjunto de dominios, cumpliéndose:

$\forall\ A1_i\ \exists\ A2_j \mid dom\ (A1_i) = dom\ (A2_j)$ y $\forall\ A2_j\ \exists\ A1_i \mid dom\ (A2_j) = dom\ (A1_i)$

es decir, cada atributo de R1 tiene su equivalente de R2 (con igual dominio) y viceversa; por tanto, el grado (número de atributos) de ambas relaciones debe ser el mismo.

También se dice que R1 y R2 son *semánticamente* equivalentes, lo que no quiere decir que los nombres de los atributos sean los mismos (esto sería sólo un aspecto sintáctico).

## Restricción ( $\sigma$ )

El operador de selección $\sigma$ aplicado a la relación R con el predicado p, denotado por:

$\sigma_p\ (R)$

produce una relación, cuyo esquema será el mismo (es decir, R) y cuya extensión r’ es:

$r'(R) = \{ t_i \in r\ (R) \mid p(t_i) \}$

p es un predicado de selección, formado por una expresión lógica integrada por cláusulas lógicas combinadas utilizando paréntesis y los operadores lógicos AND, OR y NOT.

Las cláusulas lógicas son de la forma:

$A_i \ \theta \ A_j$ o $A_i \ \theta \ c$

siendo $A_i$, $A_j$ atributos, c una constante, y $\theta$ un operador de comparación ( $>$, $<$, $=$, $\leq$, $\geq$, $\neq$).

**Proyección ( $\Pi$ )**

Sea X un subconjunto estricto y no vacío de A:

$X \subset A$ y $X \neq \phi$

la proyección de R en el contexto de X, denotada por:

$\Pi_x (R)$

es una relación r' cuyo esquema es R'(X) y cuya extensión es el conjunto de tuplas de la relación original definidas sobre los atributos X, eliminando las que resulten duplicadas:

$r'(R') = \{ t_i(X) \mid X \subset A \}$

**Unión ( $\cup$ )**

Sean dos relaciones $r_1$ y $r_2$ con esquemas R1 y R2 compatibles, la unión de ambas, denotada por:

$R1 \cup R2$

es una relación con esquema R (igual a R1 o a R2 ya que ambos son equivalentes) y con extensión r:

$r(R) = \{ t_i \mid t_i \in r_1 \lor t_i \in r_2 \}$

**Diferencia (–)**

Sean dos relaciones compatibles, con esquemas R1 y R2, la diferencia de ambas denotada por:

$R1 - R2$

es una relación con esquema R (igual a R1 o a R2 ya que ambos son equivalentes) y con extensión r:

$r(R) = \{ t_i \mid t_i \in r_1 \land t_i \notin r_2 \}$

**Producto Cartesiano (X)**

Sean las relaciones $r_1$ y $r_2$ con esquemas R1 y R2, el producto cartesiano de ambas, denotado por:

R1 x R2

es una relación de grado $n=n_1+n_2$, cuyo esquema R esta formado por los $n_1+n_2$ atributos:

$$( A1_i: D1_i, ..., A1_{n_1}: D1_{n_1}, A2_{i'}: D2_{i'}, ..., A2_{n_2}: D2_{n_2} )$$

y cuya extensión, de cardinalidad m x m', es:

$$\{ < V1_{i1}, .., V1_{in_1}, V2_{j1}, .., V2_{jn_2} > | \forall i \forall j (< V1_{i1}, .., V1_{in_1} > \in r_1 \wedge < V2_{j1}, .., V2_{jn_2} > \in r_2 )\}$$

**Expresiones del Álgebra Relacional**

Con los 5 operadores anteriores es posible construir cualquier expresión del álgebra relacional, es decir, completan el lenguaje. Los demás operadores derivados se pueden expresar en función de estos fundamentales[1].

Una expresión del AR se construye a partir de subexpresiones. Siendo R1 y R2 relaciones, p un predicado definido sobre atributos de R1, y S una lista de atributos de R1, son expresiones válidas en el AR las siguientes:

1) Una relación de la BD: R1, R2, ...
2) Una relación constante C.
3) $R1 \cup R2$ , $R1 - R2$ , R1 x R2 , $\sigma_p(R1)$ , y $\Pi_S(R1)$ .
4) Las expresiones derivadas por aplicación repetida de las tres opciones anteriores.

Dos expresiones E y E' son equivalentes si se evalúan siempre al mismo resultado sea cual sea el estado de la base de datos.

[1] Los 5 operadores fundamentales no sin fijos, pueden ser otros diferentes. Por ejemplo, en el conjunto utilizado en este documento se podría sustituir la diferencia por la intersección y el conjunto de 5 operadores resultante también completaría el lenguaje.

**Combinación (θ)**

Dadas dos relaciones, $R_1(A1_1, A1_2, ..., A1_{n1})$ y $R_2(A2_1, A2_2, A2_{n2})$, la combinación de ambas, denotada por:

$$R_1 \underset{\text{condición}}{\theta} R_2$$

es una relación de grado $n=n_1+n_2$, cuyo esquema R esta formado por los $n_1+n_2$ atributos:

$$( A1_i: D1_i, ..., A1_{n_1}: D1_{n_1}, A2_i: D2_{i'}, ..., A2_{n_2}: D2_{n_2})$$

y cuya extensión r(R), de cardinalidad $m\leq(m_1+m_2)$ es:

$$\{ < V1_{i1}, ..V1_{in_1}, V2_{j1}, ..V2_{jn_2} > |$$

$$\forall i \forall j\, (< V1_{i1}, .., V1_{in_1} > \in r_1 \wedge < V2_{j1}, .., V2_{jn_2} > \in r_2 \wedge < \text{condición} >)\}$$

siendo la condición de combinación del tipo:

$A1_i\ \theta\ A2_j$ o $(A1_{i1}\ \theta\ A2_{j1})$ and $(A1_{i2} \theta\ A2_{j2})$ ....

La operación de combinación equivale a un producto cartesiano seguido de una restricción:

$$R_1\ \theta\ R_2 \quad \equiv \quad \sigma_p(R_1 \text{ x } R_2) \quad \text{siendo p la condición}$$

**Combinación Natural (*)**

La operación de combinación natural equivale a un producto cartesiano, seguido de una restricción por igualdad y de una proyección:

$$R_1 * R_2 \quad \equiv \quad \Pi_{(A_1 \cup A_2)} \sigma_p (R_1\ x\ R_2)$$

siendo p la condición de igualdad entre todas las parejas de atributos correspondientes:

$P = (A1_{i1}=A2_{j1})$ and $(A1_{i2}=A2_{j2})$ ....

y $A_1 \cup A_2$ el conjunto de todos los atributos eliminando uno de cada pareja

**Intersección (∩)**

Dadas dos relaciones, con esquemas compatibles R1 y R2, la intersección de ambas, denotada por:

$$R1 \cap R2$$

es una relación r con esquema R (igual a R1 o a R2 ya que ambos son equivalentes) y con extensión:

$r(R) = \{ t_i / t_i \in r_1 \wedge t_i \in r_2 \}$

La intersección se puede definir en función de la diferencia:

$R1 \cap R2 = R1 - ( R1 - R2 )$

**División (:)**

Dadas dos relaciones con esquemas R1(A1) y R2(A2), de grados $n_1$ y $n_2$ respectivamente, donde $A2 \subset A1$ (y por tanto, $n_1<n_2$); la división de ambas relaciones, denotada por:

R1 : R2

es una relación r de grado $n_1 - n_2$ cuyo esquema R esta formado por los $n_1 - n_2$ atributos A1 - A2:

$(A1_1, A1_2, ..., A1_{n1}) - (A2_1, A2_2, A2_{n2})$

y cuya extensión r(R) es:

$$\left\{ < v_{i1}, ... v_{i(n_1-n_2)} > | \forall < v_{i(n_1-n_2+1)}, ... v_{in_2} > \in r_2 \ \exists < v_{i1}, ... v_{i(n_1-n_2)}, v_{i(n_1-n_2+1)}, ... v_{in} > \in r_1 \right\}$$

La división se puede expresar en función de la proyección, del producto cartesiano y de la diferencia de la siguiente forma:

$R1 : R2 = \pi_C(R1) - \pi_C (R2 \times \pi_C (R1) - R1)$

siendo C el conjunto de atributos de A menos los de B.