BASES DE DATOS
OPERADORES DEL ÁLGEBRA RELACIONAL

Definición. Llamaremos **dominios** a los siguientes conjuntos:

- $\mathbb{N}^{\mathbf{null}}$, los numeros naturales y el elemento especial **null**
- $\mathbb{R}^{\mathbf{null}}$, los numeros reales y el elemento especial **null**
- $\mathbb{S}^{\mathbf{null}}$, las cadenas y el elemento especial **null**
- y por extension a todos los posibles productos cartesianos entre ellos

Definición. Una **tupla** $(t_1, \ldots, t_n)$ sera un elemento de un dominio $\mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$.

Definición. Una **relacion** será un conjunto[1] de tuplas de un mismo dominio, que será considerado el dominio de la relacion.

Definición. Una relacion puede tener un **nombre**, y ademas un rotulo distinto para cada indice de sus tuplas. Estos rótulos seran llamados **atributos** de la relacion. La tupla ordenada de atributos de una relacion sera el **esquema de la relacion**. Si una relación no tiene esquema, diremos que tiene **esquema anónimo**. Dado un esquema $R = A_1, \ldots, A_n$ y una relacion r escribiremos $r(R)$ para **declarar** que r tiene esquema R. Utilizaremos $t[A]$ para referirnos al valor que la tupla t asocia al atributo A.

Dados dos esquemas de relacion $R = A_1, A_2, \ldots, A_n$ y $S = B_1, B_2, \ldots, B_m$ disjuntos, es decir, tal que $A_i \neq B_j$ para todo i, j, definimos la **concatenacion de esquemas** como $R, S = A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_m$.

Dada $r(R)$ donde $R = A_1, A_2, \ldots, A_n$, se define por $r.R$ al esquema de relación cuyos atributos son $r.A_1$, $r.A_2$, $\ldots$ y $r.A_n$, es decir, $r.R = r.A_1, r.A_2, \ldots, r.A_n$.

Definición. Diremos que dos relaciones r y s tienen **esquemas compatibles** si estan incluidas en el mismo dominio $\mathcal{D}$.

Ahora definiremos las funciones que utilizaremos para construir nuevas relaciones:

Union, interseccion y diferencia. Dadas dos relaciones r y s con esquemas compatibles, $r \cap s$, $r \cup s$ y $r - s$ son relaciones definidas con la nocion natural de estas operaciones de conjuntos y con esquema de relación compatible con R (y por ende, S), y anónimo.

[1] Un conjunto matematico, sin repeticiones

Producto cartesiano. Dadas $r(R)$ y $s(S)$, $r \times s$ devuelve el producto cartesiano de las relaciones r y s con esquema $r.R, s.S$ (la concatenacion de los esquemas dados). Por simplicidad notacional, suelen omitirse los prefijos "$r.$" y "$s.$" que no lleven a que dos atributos queden con el mismo rótulo. Para el caso particular $r \times r$, la relacion devuelta será de esquema anonimo.

Renombre. Dada $r(R)$, y S un esquema compatible con R, $\rho_S(r)$ devuelve la relacion r, pero con esquema de relación S en vez de R. Como *sintax sugar* podemos cambiar el nombre de solo algunos atributos. Utilizaremos $\rho_{A\to B, C\to D}(r)$ como *sintax sugar* referirnos a $\rho_{R'}(r)$ con R' al esquema resultante de reemplazar en R el atributo A es por B y el C por D.

Seleccion. Dada una condición booleana P sobre de los atributos de R formada por operaciones sobre los dominios (expresiones aritméticas, de cadenas de caracteres, de comparación de fechas, etc.), y dada una relacion $r(R)$, $\sigma_P(r)$ devuelve la relación r filtrada por la condicion P manteniendo el esquema de relación R.

Proyeccion. Dada la relacion $r(R)$ y un esquema de relación $S \subseteq R$. $\Pi_S(r)$ devuelve la relación con esquema de S y se obtiene descartando de cada tupla de r las componentes que no corresponden a atributos de S, y reordenando los restantes componentes de la tupla según el orden en que se listen los atributos en el esquema S.

Agregacion. Dada $r(R)$, ${}_A\Upsilon_{F_1(B_1),F_2(B_2),\dots,F_m(B_m)}(r)$ donde $A = A_1, \dots, A_n$ y $B = B_1, \dots, B_m$ son subesquemas de $r(R)$ y las F_j son funciones de agregación: avg (promedio), min (mínimo), max (máximo), sum (suma), count (contar) y cualquier otra funcion definida para el dominio del atributo en cuestion. Devuelve una relacion con esquema de tiene $n + m$ elementos, con los dominios correspondientes a A_i y F_j pero anónimo.

La tupla $(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) \in {}_A\Upsilon_{F_1(B_1),F_2(B_2),\dots,F_m(B_m)}(r)$ si y sólo si, considerando el conjunto $T_{x_1,\dots,x_n}$ de tuplas de t tales que $t[A_i] = x_i$ para todo i, y_j se obtiene aplicando la función F_j a los atributos B_j de todas las tuplas del conjunto $T_{x_1,\dots,x_n}$. Notar que cada tupla de la relacion resultante es la unica que comienza con$x_1, x_2, \dots, x_n$[2]

Las siguientes operaciones pueden definirse a partir de las anteriores como *sintax sugar*:

Proyección generalizada. Permite utilizar expresiones como subíndices de Π. Dada $r(R)$ se define $\Pi_{F(A)}(r) = \Upsilon_{F(A)}(r)$. Por ejemplo $\Pi_{eId, eSalario*13}(empleados) = \Upsilon_{I(eId),(\lambda x.x*13)(eSalario)}(empleados)$, devuelve un esquema anónimo.

Reunion natural. Dadas $r(R)$ y $s(S)$ con $A_1, \dots, A_n$ atributos comunes entre R y S. Definimos $r \bowtie s = \Pi_T(\sigma_{r.A_1=s.A_1 \wedge r.A_2=s.A_2 \wedge \dots \wedge r.A_n=s.A_n}(r \times s))$ donde T es el esquema R, S' y S' es S sin los atributos $A_1, \dots, A_n$. El esquema de $r \bowtie s$ es T.

[2]En otras palabras, $A_1, A_2, \dots, A_n$ es superclave de la relacion devuelta.

Reunion externa. Las filminas muestran cómo las **reuniones externas** pueden obtenerse a partir de la reunión natural y otros operadores.

División. Sean $r(R)$ y $s(S)$ con S tramo final de R. Sea T el tramo inicial de R sin los atributos de S. La **división** entre r y s se define por $r \div s = \Pi_T(r) - \Pi_T((\Pi_T(r) \times s) - r)$. Su esquema es T.