

## Capítulo 6

# El Álgebra Relacional

### 6.1. Introducción

Como ya hemos mencionado en el capítulo anterior, el Álgebra y el Cálculo Relacional son los dos formalismos o herramientas que se desarrollaron para modelar el proceso de obtención de información a partir de una base de datos relacional. Aunque utilizan principios diferentes, ambos persiguen la misma finalidad: obtener, como resultado de una pregunta a una base de datos relacional, una relación adecuada. La diferencia entre ambos formalismos se encuentra en el enfoque que se utiliza para la obtención de dicha relación:

- En el álgebra relacional, se sigue un sistema “procedimental”, obteniendo el resultado mediante la aplicación de sucesivas operaciones sobre las relaciones de la base de datos.
- En el cálculo relacional, se define el resultado como el conjunto de valores que hacen cierta una determinada fórmula lógica. Este último enfoque queda fuera del ámbito de este texto, por lo que nos centraremos en el Álgebra Relacional (AR, a partir de ahora). Para un estudio detallado del cálculo relacional véase [10, 8]

El AR fue introducida por Codd en el mismo trabajo donde se estableció el modelo relacional. Inicialmente, se definieron ocho operadores, que luego fueron ampliados por diversos autores. No obstante, como ya fue probado por Codd en 1974, los ocho operadores iniciales del AR son suficientes para reproducir cualquier consulta

formulada en cálculo relacional, lo que implica que estos ocho operadores tienen la misma capacidad expresiva que la lógica de primer orden.

Operador	Notación
Selección	$\sigma$
Proyección	$\pi$
Unión	$\cup$
Intersección	$\cap$
Diferencia	$-$
Producto cartesiano	$\times$
$\Theta$ -Reunión	$\bowtie_{\Theta}$
División	$\div$

Tabla 6.1: Operadores del Álgebra Relacional

Estos ocho operadores se muestran en la tabla 6.1 y pueden clasificarse con respecto a distintos criterios.

- Con respecto al tipo de operador:

**Operadores monarios:** Son aquellos que se aplican a una relación para dar lugar a otra, concretamente, los operadores de *selección* y *proyección*.

**Operadores binarios:** Son aquellos que se aplican a dos relaciones para dar lugar a una tercera, concretamente, los operadores de *unión*, *intersección*, *diferencia*, *producto cartesiano*,  *$\theta$ -reunión* y *división*.

- Con respecto a su relación con el modelo relacional:

**Operadores conjuntistas:** Se basan en el hecho de que las relaciones son conjuntos de tuplas: los operadores de *unión*, *diferencia*, *intersección* y *producto cartesiano*.

**Operadores relacionales:** Son más específicos del concepto de relación: los operadores de *selección*, *proyección*,  *$\Theta$ -reunión* y *división*.

También hay que indicar que los ocho operadores no son independientes y que, como veremos, se pueden expresar unos en función de otros. Hay cinco operadores fundamentales, siendo posible obtener la  $\theta$ -reunión, la intersección y división en función de ellos.

Dedicaremos el resto del tema al estudio de estos ocho operadores, insistiendo en sus distintos usos para resolver consultas. También estudiaremos las extensiones más interesantes de los mismos.

Antes de comenzar la descripción detallada de los operadores y con objeto de que el lector tenga toda la información necesaria para seguir los ejemplos que vamos a desarrollar a lo largo de este capítulo, recordemos el esquema de la base de datos que se va a utilizar en los ejemplos. Dicha base de datos ya fue descrita en capítulos anteriores y su esquema es el siguiente:

ALUMNOS	DNI	NOM_ALUM	EDAD	DIRECCION	PROVINCIA	BECA
ASIGNATURAS	COD_ASIG	NOM_ASIG	CREDITOS	CARACTER	CURSO	
PROFESORES	NRP	NOM_PROF	CATEGORIA	AREA	COD_DEP	
DEPARTAMENTOS	COD_DEP	NOM_DEP	DIRECTOR			
AULAS	COD-AULA	CAPACIDAD				
GRUPOS	COD_GRUP	COD_ASIG	TIPO	NRP	MAX_AL	
CLASE	DIA	HORA	COD_AULA	COD_GRUP	COD_ASIG	TIPO
MATRICULAS	COD_ASIG	COD_GRUP	TIPO	DNI	CALIFICACION	CONVOCATORIA

Figura 6.1: Esquema de la base de datos utilizada en los ejemplos

Si bien el esquema es el mismo que el obtenido en el capítulo 4, se han cambiado los nombres a algunos atributos y hemos añadido alguno nuevo para clarificar las consultas que utilizemos como ejemplo.

## 6.2. Operadores relacionales monarios

Estos dos operadores, selección y proyección, nos van a permitir obtener cualquier parte de una tabla y, por ello, se denominan de forma coloquial “operadores tijera”. Antes de introducirlos vamos a precisar un poco la notación que emplearemos en este capítulo. Sea  $R[A_1, \dots, A_n]$  una relación cualquiera:

- Notaremos por  $P(A_i, A_j)$  a cualquier expresión lógica del tipo  $A_i \theta A_j$ , donde  $\theta$  es un operador de comparación ( $=, >, <, \leq, \neq, \geq$ ). Obviamente,  $A_j$  se puede sustituir por una constante.

- Notaremos por  $\Theta(A_1, \dots, A_l)$  a toda propiedad lógica asociada al conjunto de atributos  $\{A_1 \dots A_l\}$  que sea combinación mediante  $\{\vee, \wedge, \neg\}$  de propiedades atómicas, incluyendo constantes y nombres de atributos de  $\{A_1 \dots A_l\}$ .

### 6.2.1. La operación de selección

Esta operación es la más elemental que podemos suponer asociada a una tabla y lo que hace es simplemente seleccionar (valga la redundancia) un conjunto de filas de una tabla dada. Formalmente, se define de la siguiente manera:

**Definición 6.1 (Selección).** Sea  $R[A_1, \dots, A_n]$  una relación cualquiera,  $\Theta$  una propiedad asociada a  $\{A_1 \dots A_n\}$  y  $r$  una instancia de  $R$ ; el operador  $\Theta$ -selección aplicado a  $r$  y que notaremos por  $\sigma_{\Theta}(r)$  obtiene aquellas tuplas de  $r$  para las que  $\Theta$  es cierta.

NRP	NOM_PROF	CATEGORIA	AREA	COD_DEP
2428456	Juan Sanchez Perez	AS	COMPUT	CCIA
24283256	Antonia Perez Rodriguez	CU	COMPUT	CCIA
242256	Luis Perez Perez	TE	LENGUA	LSI
84256	Carmen Perez Sanchez	TU	LENGUA	LSI
324256	David Perez Jimenez	CU	ARQUIT	ATC
24256	Maria Lopez Ruiz	TU	ARQUIT	ATC
2842560	Jose Alvarez Perez	CE	ELECTR	ELEC
842560	Adela Perez Sanchez	AS	ELECTR	ELEC
84560	Luis Martinez Perez	AS	TSEÑAL	TESE
242560	Maria Gomez Sanchez	CU	TSEÑAL	TESE

Tabla 6.2: Instancia inicial de la relación Profesores

**Ejemplo 6.1.** Consideremos la instancia de la tabla Profesores que aparece en la tabla 6.2, el resultado de la selección

$$\sigma_{\text{categoria}='AS'}(\text{profesores})$$

aparece en la tabla 6.3 y nos da aquellos profesores con categoría AS.

NRP	NOM_PROF	CATEGORIA	AREA	COD_DEP
2428456	Juan Sanchez Perez	AS	COMPUT	CCIA
842560	Adela Perez Sanchez	AS	ELECTR	ELEC
84560	Luis Martinez Perez	AS	TSEÑAL	TESE

Tabla 6.3: Resultado del ejemplo 6.1.

**Ejemplo 6.2.** *El resultado de la selección*

$$\sigma_{\text{categoria} \neq 'AS' \wedge (\text{area} = 'COMPUT' \vee \text{area} = 'ELECTR')}(profesores)$$

aparece en la tabla 6.4 y nos da aquellos profesores cuya categoría no es AS y cuya área de conocimiento es COMPUT o ELECTR.

NRP	NOM_PROF	CATEGORIA	AREA	COD_DEP
24283256	Antonia Perez Rodriguez	CU	COMPUT	CCIA
2842560	Jose Alvarez Perez	CE	ELECTR	ELEC

Tabla 6.4: Resultado del ejemplo 6.2.

**6.2.2. La operación de proyección**

Si la selección puede verse como un “corte horizontal” en una tabla, la proyección se puede considerar intuitivamente como un “corte vertical”. Si la primera nos permite obtener las filas que nos interesan de una tabla, la segunda nos permite elegir las columnas que necesitamos.

**Definición 6.2 (Proyección).** Sea  $R[A_1..A_n]$  una relación cualquiera,  $\{A_i...A_j\}$  un subconjunto de sus atributos y  $r$  una instancia de  $R$ ; el operador proyección sobre  $\{A_i...A_j\}$  aplicado a  $r$ , que notaremos por  $\pi_{A_i...A_j}(r)$ , obtiene tuplas de  $r$  eliminando de la tabla aquellos atributos no pertenecientes a  $\{A_i...A_j\}$  y quitando las tuplas redundantes.

**Ejemplo 6.3.** Consideremos la tabla profesores que aparece en la tabla 6.2. El resultado de

$$\pi_{NRP, nom\_prof, categoria}(Profesores)$$

aparece en la tabla 6.5.

NRP	NOM.PROF	CATEGORIA
2428456	Juan Sanchez Perez	AS
24283256	Antonia Perez Rodriguez	CU
242256	Luis Perez Perez	TE
84256	Carmen Perez Sanchez	TU
324256	David Perez Jimenez	CU
24256	Maria Lopez Ruiz	TU
2842560	Jose Alvarez Perez	CE
842560	Adela Perez Sanchez	AS
84560	Luis Martinez Perez	AS
242560	Maria Gomez Sanchez	CU

Tabla 6.5: Resultado del ejemplo 6.3.

Es importante tener en cuenta el detalle de la eliminación de tuplas redundantes. En una tabla no puede haber tuplas repetidas y, por tanto, cuando en el resultado de una proyección aparecen tuplas repetidas, estas se deben descartar. Si en el conjunto de atributos sobre el que se proyecta no se incluye una llave candidata, la tabla resultado puede incluir tuplas repetidas que el operador proyección elimina. El siguiente ejemplo aclara algo más este punto:

**Ejemplo 6.4.** *La clave primaria de la tabla Profesores es NRP, y nom\_prof es una clave candidata. Si proyectamos la tabla sobre {area,cod\_dep}, atributos que describen el área de conocimiento y el código del departamento al que pertenecen los profesores, pero no eliminamos las tuplas repetidas obtenemos la tabla 6.6. El verdadero resultado de la operación de proyección, obtenido tras la eliminación de tuplas repetidas, aparece en la tabla 6.7.*

Cuando queramos saber cuál es el dominio “activo” de un atributo cualquiera, es decir, obtener *todos* los valores que toma dicho atributo en una instancia dada de la BD, sólo tenemos que tomar la proyección de ese atributo sobre la tabla correspondiente.

**Ejemplo 6.5.** *El dominio activo del atributo categoría se obtiene calculando*

$$\pi_{categoría}(profesores)$$

*y aparece en la tabla 6.8.*

AREA	COD_DEP
COMPUT	CCIA
COMPUT	CCIA
LENGUA	LSI
LENGUA	LSI
ARQUIT	ATC
ARQUIT	ATC
ELECTR	ELEC
ELECTR	ELEC
TSEÑAL	TESE
TSEÑAL	TESE

Tabla 6.6: Resultado sin eliminar tuplas.

AREA	COD_DEP
COMPUT	CCIA
LENGUA	LSI
ARQUIT	ATC
ELECTR	ELEC
TSEÑAL	TESE

Tabla 6.7: Eliminación de tuplas redundantes.

### 6.2.3. Composición de operadores

Como comentamos en la introducción del tema, el AR se basa en la aplicación sucesiva de operadores hasta que obtenemos la tabla solución a nuestra consulta. Esta forma de actuar se pone de manifiesto si utilizamos sobre una tabla los dos operadores que conocemos, ya que aplicando el operador proyección al resultado de una selección podemos obtener la parte que deseemos de una tabla.

**Ejemplo 6.6.** Si deseamos saber el NRP y el Nombre de aquellos profesores que pertenecen al departamento cuyo código es ELEC, la forma de obtenerlo es:

1. Seleccionamos los profesores con departamento igual a ELEC, es decir:

$$\sigma_{cod\_dep=ELEC}(profesores)$$

La solución de esta selección aparece en la tabla 6.9.

CATEGORIA
AS
CU
TE
TU
CE

Tabla 6.8: Resultado del ejemplo 6.5.

2. Aplicamos al resultado anterior la proyección sobre los atributos que deseamos; es decir, la expresión completa sería:

$$\pi_{NRP, nom\_prof}(\sigma_{cod\_dep=ELEC}(profesores))$$

La solución final aparece en la tabla 6.10.

NRP	NOM_PROF	CATEGORIA	AREA	COD_DEP
2842560	Jose Alvarez Perez	CE	ELECTR	ELEC
842560	Adela Perez Sanchez	AS	ELECTR	ELEC

Tabla 6.9: Resultado de la primera consulta del ejemplo 6.6.

NRP	NOM_PROF
2842560	Jose Alvarez Perez
842560	Adela Perez Sanchez

Tabla 6.10: Resultado de la segunda consulta del ejemplo 6.6.

A continuación, proponemos algunos ejemplos de composición de operadores, algunos referidos a la tabla PROFESORES y otros referentes a otras tablas de la base de datos de la que forma parte:

**Ejemplo 6.7.** Encontrar los nombres de los profesores que no tienen categoría AS y pertenecen a las áreas de conocimiento TSEÑAL o ARQUIT.

Resultado:

$$\pi_{nom\_prof}(\sigma_{(categoria \neq AS) \wedge (area=TSEÑAL \vee area=ARQUIT)}(profesores))$$



**Ejemplo 6.8.** *Encontrar las áreas de conocimiento que tienen profesores con categoría CU o TU.*

Resultado:

$$\pi_{area}(\sigma_{(categoria=TU \vee categoria=CU)}(profesores))$$

**Ejemplo 6.9.** *Encontrar el nombre y la edad de aquellos alumnos que vienen de GRANADA.*

Resultado:

$$\pi_{nom\_alum, edad}(\sigma_{provincia=GRANADA}(alumnos))$$

**Ejemplo 6.10.** *Encontrar el DNI y el nombre de aquellos alumnos que tienen más de 25 años.*

Resultado:

$$\pi_{DNI, nom\_alum}(\sigma_{edad \geq 25}(alumnos))$$

**Ejemplo 6.11.** *Encontrar las provincias de las que vienen alumnos becados.*

Resultado:

$$\pi_{provincia}(\sigma_{beca=SI}(alumnos))$$

## 6.3. El producto cartesiano y la $\Theta$ -reunión

Vamos ahora a estudiar los operadores que permiten “conectar” tablas. De manera informal, se puede decir que, si la selección y proyección son operadores “tijera”, la  $\Theta$ -reunión es un operador “pegamento”. Como ya comentamos, este es un operador derivado que se define a partir de otros dos más, concretamente, a partir del producto cartesiano y de la selección. Centrémonos primero en el producto cartesiano.

### 6.3.1. El operador producto cartesiano

El producto cartesiano es un operador binario que genera todas las tuplas que se pueden obtener combinando una tupla de la primera relación con otra tupla de la segunda relación.