

Tema XI

Conceptos a abordar en este material

- Álgebra Relacional
 - Operaciones del Modelo Relacional.
-

Álgebra Relacional

El álgebra relacional y los cálculos relacionales, son dos lenguajes formales del **modelo relacional** que sirve para definir un conjunto de operaciones básicas para manipular la base de datos. Es decir, el modelo de datos incluye tanto la definición de la estructura de las relaciones y las restricciones, y además un conjunto de operaciones que se pueden realizar sobre los datos. Esas operaciones las escribimos en un lenguaje formal que llamamos álgebra relacional.

Cada operación que escribimos usando álgebra relacional, actúa sobre los datos generando resultados que son nuevas relaciones y esas operaciones sirven como base para la implementación de consultas en los DBMS (Administradores de Bases de Datos).

Las operaciones del álgebra relacional se dividen en dos grandes grupos, uno tiene que ver con el conjunto de operaciones de la teoría matemática de conjuntos (UNION, INTERSECTION, SET DIFFERENCE, CARTESIAN PRODUCT) y el otro son operaciones desarrolladas específicamente para las bases de datos relaciones (SELECT, PROJEC, JOIN, etc).

Segundo grupo: Operaciones específicas de bases de datos relacionales.**Selección**

Esta operación es unaria, porque actúa sobre una sola relación. Lo que hace es seleccionar un subconjunto de tuplas, que satisfacen una condición. La sintaxis es la siguiente

$\sigma_{\langle \text{condición de selección} \rangle (R)}$

Ejemplo: $\sigma_{\text{CodTratamiento}=12} (\text{VISITA})$

$\sigma_{\text{Fecha}<30.09.2003.08.00.23} (\text{VISITA})$

El símbolo σ especifica la operación de selección, luego la expresión lógica se especifica sobre los atributos de la relación, y en mayúsculas escribimos el nombre de la relación sobre la que se aplica la operación. Es importante destacar que el resultado de la operación será una nueva relación con los mismos atributos que la relación sobre la que se aplica originalmente.

En el primer ejemplo estamos seleccionando todas las tuplas de la relación VISITA que cumplen con que en el atributo CodTratamiento el valor es 12.

En el segundo ejemplo se seleccionan todas las tuplas de la relación VISITA que cumplen con que el atributo Fecha es menor que el día 30 del mes 09 del año 2003....

Con respecto a la expresión lógica debemos respetar la siguiente forma: (nombre del atributo, operador de comparación, valor constante) o (nombre del atributo, operador de comparación, nombre del atributo). Donde operador de comparación puede ser alguno de los siguientes: =, <, ≤, >, ≥, ≠. y las condiciones de selección se pueden relacionar mediante operadores lógicos and, or, not. Cuando se trata de atributos cuyos dominios son valores ordenados, sí se puede usar los operadores de comparación menor que o mayor igual que, pero cuando el dominio de valores no es ordenado sólo podemos usar igual o diferente que. A continuación se muestra un ejemplo de atributos con dominios de valores ordenados.

Ejemplo:

$\sigma_{(\text{Dno}=4 \text{ AND } \text{Sueldo}>25000) \text{ OR } (\text{Dno}=5 \text{ AND } \text{Sueldo}>30000)} (\text{EMPLEADO})$

Nombre	Apellido1	Apellido2	Dni	FechaNac	Dirección	Sexo	Sueldo	SuperDni	Dno
Alberto	Campos	Sastre	333445555	08-12-1955	Avda. Ríos, 9	H	40000	888665555	5
Juana	Sainz	Oreja	987654321	20-06-1941	Cerquillas, 67	M	43000	888665555	4
Fernando	Ojeda	Ordóñez	666884444	15-09-1962	Portillo, s/n	H	38000	333445555	5

Si el dominio de valores es desordenado, tendríamos por ejemplo esta situación:

$\sigma_{(\text{Nombre}=\{\text{Juana, Ana, Valentina}\})} (\text{EMPLEADO})$

Esto selecciona todas las tuplas de la relación Empleado, que cumplan con que el nombre del empleado sea Juana, o Ana o Valentina.

Proyección

También es una operación unaria, actúa sobre una sola relación, y lo que hace es seleccionar un subconjunto de atributos, o columnas que cumple con la lista de atributos que indicamos. Puede verse también como una partición vertical de una tabla, que queda en dos partes, una que tiene las columnas que nos interesan, y otra que tiene las columnas que se descartan. La sintaxis es la siguiente :

$$\pi_{\langle \text{lista de atributos} \rangle}(R)$$

Por ejemplo: $\pi_{\langle \text{Apellido1, Nombre, Sueldo} \rangle}(\text{EMPLEADO})$ obtiene una lista de los apellidos, nombres y sueldos de todos los empleados.

Apellido1	Nombre	Sueldo
Pérez	José	30000
Campos	Alberto	40000
Jiménez	Alicia	25000
Sainz	Juana	43000
Ojeda	Fernando	38000
Oliva	Aurora	25000
Pajares	Luis	25000
Ochoa	Eduardo	55000

Es importante destacar que la operación de proyección elimina de su resultado las tuplas duplicadas, esto pasa cuando hacemos una proyección en cuya lista de atributos no incluimos los que o el que es clave primaria, por lo tanto las tuplas se pueden repetir. **La proyección sólo nos muestra las tuplas únicas, si hay repetidas no las muestra.**

Primer grupo: Operaciones del álgebra relacional de las teoría de conjuntos.

Unión

Esta operación se aplica sobre dos relaciones o conjuntos, donde debemos asegurarnos que ambos tengan la misma cantidad de atributos y que se correspondan sus dominios. El resultado de esta operación incluye todas las tuplas de un conjunto y todas las tuplas del otro y cuando una tupla esté en ambos conjuntos también formará parte del resultado, luego si una tupla aparece dos veces, formará parte del resultado una sola ocurrencia. Veamos un ejemplo: Si queremos tener una lista de

los DNI de los empleados que trabajan en el departamento 5 o que supervisan a éstos, podemos hacer lo siguiente.

Primero seleccionamos todos los empleados que trabajan en el departamento 5 de la tabla Empleado. $\sigma_{(Dno=5)}(EMPLEADO)$

Luego me quedo sólo con una de las columna que me interesan de esa nueva tabla que es Dni de los empleados $\pi_{<Dni>}(\sigma_{(Dno=5)}(EMPLEADO))$

Por otra parte me quedo con la otra columna que me interesa de esa nueva tabla que es DniSuper $\pi_{<DniSuper>}(\sigma_{(Dno=5)}(EMPLEADO))$

Y finalmente uno ambas tablas para obtener lo que busco que es el dni de todos los empleados del departamento 5 o quienes los supervisan.

$\pi_{<Dni>}(\sigma_{(Dno=5)}(EMPLEADO)) \cup \pi_{<DniSuper>}(\sigma_{(Dno=5)}(EMPLEADO))$

Dni
123456789
333445555
666884444
453453453

Dni
333445555
888665555

Dni
123456789
333445555
666884444
453453453
888665555

Intersección

El resultado de esta operación incluye todas las tuplas que tengan en común los dos conjuntos. Es una operación conmutativa que se puede aplicar a dos o más relaciones a la vez. Ejemplo: Si queremos ver el nombre y apellido de todos los estudiantes que también son profesores escribimos:

ESTUDIANTE \cap PROFESOR

ESTUDIANTE

Nombre	Apellido
Susana	Gómez
Luis	Campos
Juan	Garrido
Bárbara	Durán
Amanda	González
Joaquín	Martín
Ernesto	Flores

PROFESOR

Nom	Apell
Antonio	Fernández
Ricardo	Adriano
Susana	Gómez
Francisco	Peláez
Luis	Campos

Nombre	Apellido
Susana	Gómez
Luis	Campos

Diferencia de conjuntos

Esta operación no es conmutativa y se aplica siempre a dos conjuntos. Su resultado es una nueva relación que incluye todas las tuplas que están en el primer conjunto pero no en el segundo, es decir todas las tuplas del primer conjunto quitando de él las tuplas que coincidan con el segundo conjunto. Ejemplo: Si queremos obtener una lista de los nombres y apellidos de las personas que sólo son profesores pero no estudiantes tendríamos la siguiente expresión.

PROFESOR - ESTUDIANTE

Nom	Apell
Antonio	Fernández
Ricardo	Adriano
Francisco	Peláez

Producto Cartesiano

También conocido como producto cruzado o cross join, se trata de una operación de dos conjuntos que no necesariamente deben compartir cantidad de atributos y dominio de los mismos. El resultado será una nueva relación que combinará cada tupla de un conjunto con cada tupla del otro. La cantidad de tuplas de la nueva relación se obtiene multiplicando la cantidad de tuplas de las dos relaciones que se estén asociando.

Veámos este ejemplo: Queremos tener una lista de los empleados subordinados a una empleada femenina. Partiendo de las siguientes relaciones:

EMPLEADO

Nombre	Apellido1	Apellido2	Dni	FechaNac	Dirección	Sexo	Sueldo	SuperDni	Dno
José	Pérez	Pérez	123456789	01-09-1965	Eloy I, 98	H	30000	333445555	5
Alberto	Campos	Sastre	333445555	08-12-1955	Avda. Ríos, 9	H	40000	888665555	5
Alicia	Jiménez	Celaya	999887777	12-05-1968	Gran Vía, 38	M	25000	987654321	4
Juana	Sainz	Oreja	987654321	20-06-1941	Cerquillas, 67	M	43000	888665555	4
Fernando	Ojeda	Ordóñez	666884444	15-09-1962	Portillo, s/n	H	38000	333445555	5
Aurora	Oliva	Avezuela	453453453	31-07-1972	Antón, 6	M	25000	333445555	5
Luis	Pajares	Morera	987987987	29-03-1969	Enebro, 90	H	25000	987654321	4
Eduardo	Ochoa	Paredes	888665555	10-11-1937	Las Peñas, 1	H	55000	NULL	1

SUBORDINADO

DniEmpleado	NombSubordinado	Sexo	FechaNac	Relación
333445555	Alicia	M	05-04-1986	Hija
333445555	Teodoro	H	25-10-1983	Hijo
333445555	Luisa	M	03-05-1958	Esposa
987654321	Alfonso	H	28-02-1942	Esposo
123456789	Miguel	H	04-01-1988	Hijo
123456789	Alicia	M	30-12-1988	Hija
123456789	Elisa	M	05-05-1967	Esposa

Lo primero que hacemos es seleccionar las tuplas de acuerdo al sexo y proyectar sólo los atributos de nombre, apellido y dni de las empleadas, lo que expresamos de la siguiente forma:

$\pi_{\langle \text{Nombre, Apellido1, Dni} \rangle}(\sigma_{(\text{Sexo}=\text{M})}(\text{EMPLEADO}))$ obteniendo el siguiente resultado

NOMBRES_EMPLEADOS

Nombre	Apellido1	Dni
Alicia	Jiménez	999887777
Juana	Sainz	987654321
Aurora	Oliva	453453453

Si realizo la operación Producto Cruzado entre la relación Nombres_Empleados y Subordinados, vamos a tener 21 tuplas sin sentido, porque se van a vincular cada tupla de Nombres_Empleados con cada tupla de Subordinado, cuando sólo tiene coherencia que el DniEmpleado de la relación Subordinado coincida con el Dni de la relación Nombres_Empleados. Veámos el resultado absurdo de la operación Nombres_Empleados \times Subordinados.

EMPLEADOS_SUBORDINADOS

Nombre	Apellido1	Dni	DniEmpleado	NombreSubordinado	Sexo	FechaNac	...
Alicia	Jiménez	999887777	333445555	Ana	M	05-04-1986	...
Alicia	Jiménez	999887777	333445555	Teodoro	H	25-10-1983	...
Alicia	Jiménez	999887777	333445555	Ruth	M	03-05-1958	...
Alicia	Jiménez	999887777	987654321	Augusto	H	28-02-1942	...
Alicia	Jiménez	999887777	123456789	Miguel	H	01-04-1988	...
Alicia	Jiménez	999887777	123456789	Ana	M	30-12-1988	...
Alicia	Jiménez	999887777	123456789	Elisa	M	05-05-1967	...
Juana	Sainz	987654321	333445555	Ana	M	05-04-1986	...
Juana	Sainz	987654321	333445555	Teodoro	H	25-10-1983	...
Juana	Sainz	987654321	333445555	Ruth	M	03-05-1958	...
Juana	Sainz	987654321	987654321	Augusto	H	28-02-1942	...
Juana	Sainz	987654321	123456789	Miguel	H	04-01-1988	...
Juana	Sainz	987654321	123456789	Ana	M	30-12-1988	...
Juana	Sainz	987654321	123456789	Elisa	M	05-05-1967	...
Aurora	Oliva	453453453	333445555	Ana	M	05-04-1986	...
Aurora	Oliva	453453453	333445555	Teodoro	H	25-10-1983	...
Aurora	Oliva	453453453	333445555	Ruth	M	03-05-1958	...
Aurora	Oliva	453453453	987654321	Augusto	H	28-02-1942	...
Aurora	Oliva	453453453	123456789	Miguel	H	04-01-1988	...
Aurora	Oliva	453453453	123456789	Ana	M	30-12-1988	...
Aurora	Oliva	453453453	123456789	Elisa	M	05-05-1967	...

Las únicas tuplas que tienen sentido de esta relación son aquellas en que coincide el Dni con el DniEmpleado, el resto no. Por lo tanto para lograr que se muestren sólo las tuplas que mantienen la coherencia entre los datos debemos realizar una selección de las tuplas de esta tabla que cumplan la condición de que el Dni sea igual al DniEmpleado. Al final la operación completa nos queda como sigue: ($\sigma_{(Dni=DniEmpleado)}(NOMBRES_EMPLEADOS \times SUBORDINADOS)$).

Como es tan común realizar este tipo de operaciones entre relaciones, lo que se hace es recurrir a una operación llamada Concatenación o Join que integra la selección y el producto cruzado.

Concatenación - Join

Esta operación se utiliza para combinar tuplas relacionadas de dos relaciones en una sola. Se utiliza de la siguiente forma (siguiendo el ejemplo anterior):

NOMBRES_EMPLEADOS $\bowtie_{\text{Dni}=\text{DniEmpleado}}$ SUBORDINADOS.

En forma genérica:

$R \bowtie_{\langle \text{Condición de Conexión} \rangle} S$

En este tipo de operaciones las claves foráneas juegan un papel fundamental, ya que señalan los atributos que se relacionan entre dos relaciones.

La principal diferencia entre el producto cartesiano y la concatenación es la cantidad de tuplas que se tienen como resultado de cada operación. En la concatenación sólo aparecen las tuplas que satisfacen la conexión, mientras que en el producto cartesiano aparecen todas las tuplas (las de una relación multiplicadas por las de la otra relación). Recordar que la condición de conexión puede estar relacionada con otra condición mediante operadores lógicos y cada condición en sí debe contener operadores relacionales.

Claves foráneas de la base de datos relacional Empresa

EMPLEADO

Nombre	Apellido1	Apellido2	<u>Dni</u>	FechaNac	Dirección	Sexo	Sueldo	SuperDni	Dno
--------	-----------	-----------	------------	----------	-----------	------	--------	----------	-----

DEPARTAMENTO

NombreDpto	<u>NumeroDpto</u>	DniDirector	FechaIngresoDirector
------------	-------------------	-------------	----------------------

LOCALIZACIONES_DPTO

<u>NumeroDpto</u>	<u>UbicacionDpto</u>
-------------------	----------------------

PROYECTO

NombreProyecto	<u>NumProyecto</u>	UbicacionProyecto	NumDptoProyecto
----------------	--------------------	-------------------	-----------------

TRABAJA_EN

<u>DniEmpleado</u>	<u>NumProy</u>	Horas
--------------------	----------------	-------

SUBORDINADO

<u>DniEmpleado</u>	<u>NombSubordinado</u>	Sexo	FechaNac	Relación
--------------------	------------------------	------	----------	----------

Esquema de la base de datos relacional EMPRESA.**EMPLEADO**

Nombre	Apellido1	Apellido2	<u>Dni</u>	FechaNac	Dirección	Sexo	Sueldo	SuperDni	Dno
--------	-----------	-----------	------------	----------	-----------	------	--------	----------	-----

DEPARTAMENTO

NombreDpto	<u>NumeroDpto</u>	DniDirector	FechaIngresoDirector
------------	-------------------	-------------	----------------------

LOCALIZACIONES_DPTO

<u>NumeroDpto</u>	<u>UbicacionDpto</u>
-------------------	----------------------

PROYECTO

NombreProyecto	<u>NumProyecto</u>	UbicacionProyecto	NumDptoProyecto
----------------	--------------------	-------------------	-----------------

TRABAJA_EN

<u>DniEmpleado</u>	<u>NumProy</u>	Horas
--------------------	----------------	-------

SUBORDINADO

<u>DniEmpleado</u>	<u>NombSubordinado</u>	Sexo	FechaNac	Relación
--------------------	------------------------	------	----------	----------

Estado de la base de datos relacional EMPRESA**EMPLEADO**

Nombre	Apellido1	Apellido2	Dni	FechaNac	Dirección	Sexo	Sueldo	SuperDni	Dno
José	Pérez	Pérez	123456789	01-09-1965	Eloy I, 98	H	30000	333445555	5
Alberto	Campos	Sastre	333445555	08-12-1955	Avda. Ríos, 9	H	40000	888665555	5
Alicia	Jiménez	Celaya	999887777	12-05-1968	Gran Vía, 38	M	25000	987654321	4
Juana	Sainz	Oreja	987654321	20-06-1941	Cerquillas, 67	M	43000	888665555	4
Fernando	Ojeda	Ordóñez	666884444	15-09-1962	Portillo, s/n	H	38000	333445555	5
Aurora	Oliva	Avezuela	453453453	31-07-1972	Antón, 6	M	25000	333445555	5
Luis	Pajares	Morera	987987987	29-03-1969	Enebro, 90	H	25000	987654321	4
Eduardo	Ochoa	Paredes	888665555	10-11-1937	Las Peñas, 1	H	55000	NULL	1

DEPARTAMENTO

NombreDpto	NumeroDpto	DniDirector	FechaIngresoDirector
Investigación	5	333445555	22-05-1988
Administración	4	987654321	01-01-1995
Sede Central	1	888665555	19-06-1981

LOCALIZACIONES_DPTO

NumeroDpto	UbicacionDpto
1	Madrid
4	Gijón
5	Valencia
5	Sevilla
5	Madrid

PROYECTO

NombreProyecto	NumProyecto	UbicacionProyecto	NumDptoProyecto
ProductoX	1	Valencia	5
ProductoY	2	Sevilla	5
ProductoZ	3	Madrid	5
Computación	10	Gijón	4
Reorganización	20	Madrid	1
Comunicaciones	30	Gijón	4

TRABAJA_EN

DniEmpleado	NumProy	Horas
123456789	1	32,5
123456789	2	7,5
666884444	3	40,0
453453453	1	20,0
453453453	2	20,0
333445555	2	10,0
333445555	3	10,0
333445555	10	10,0
333445555	20	10,0
999887777	30	30,0
999887777	10	10,0
987987987	10	35,0
987987987	30	5,0
987654321	30	20,0
987654321	20	15,0
888665555	20	NULL

SUBORDINADO

DniEmpleado	NombSubordinado	Sexo	FechaNac	Relación
333445555	Alicia	M	05-04-1986	Hija
333445555	Teodoro	H	25-10-1983	Hijo
333445555	Luisa	M	03-05-1958	Esposa
987654321	Alfonso	H	28-02-1942	Esposo
123456789	Miguel	H	04-01-1988	Hijo
123456789	Alicia	M	30-12-1988	Hija
123456789	Elisa	M	05-05-1967	Esposa