CAPÍTULO 5°: DEPENDENCIAS FUNCIONALES Y ÁLGEBRA RELACIONAL

1 ***Dependencias funcionales*** .-

Las dependencias funcionales determinan una manera de definir restricciones en un esquema relacional.

Dada una relación **R** que contiene los atributos **X** e **Y** se dice que **Y** *depende funcionalmente de* **X** ($X \rightarrow Y$) sí y sólo sí en todo momento cada valor de **X** tiene asociado un solo valor de **Y**. Esto es lo mismo que decir que si dos tuplas de **R** tienen el mismo valor para su atributo **X** forzosamente han de tener el mismo valor para el atributo **Y**.

A veces, para determinar el valor de un atributo es preciso conocer el valor de varios atributos, así puede ser que si el atributo **Y** no depende funcionalmente del atributo **X** ($X \not\rightarrow Y$) y simultáneamente el atributo **Y** tampoco depende funcionalmente del atributo **Z** ($Z \not\rightarrow Y$) sin embargo **Y** dependa funcionalmente de la composición de los atributos **X** y **Z** ($X \bullet Z \rightarrow Y$).

Dependencias completa y parcial: Dado un atributo compuesto **X** formado por los atributos X_1 y X_2 se dice que el atributo **Y** tiene una *dependencia funcional completa* con respecto a **X** sí:

$$X_1 \not\rightarrow Y \; ; \; X_2 \not\rightarrow Y \; ; \; Y \not\rightarrow X \qquad \text{pero} \quad X \rightarrow Y$$

Por el contrario, dado un atributo compuesto **X** formado por los atributos X_1 y X_2 se dice que el atributo **Y** tiene una *dependencia funcional parcial* con respecto a **X** si:

$$X_1 \rightarrow Y \quad \text{o} \quad X_2 \rightarrow Y \quad \text{y} \quad X \rightarrow Y$$

Axiomas de Armstrong: Son un conjunto de reglas que permiten deducir implicaciones lógicas en un conjunto de dependencias funcionales.

Si se considera una tabla relacional o una relación **T**, de la cual son atributos **X, Y, Z, W**:

Reflexividad: Si los valores del conjunto de atributos **Y** están incluidos o son iguales a un conjunto de atributos **X** se verifica que **Y** depende funcionalmente de **X**:

$$\text{Si } Y \subset X \Rightarrow X \rightarrow Y$$

Aumentatividad: Si el conjunto de atributos **Y** depende funcionalmente de **X**, la dependencia se mantiene si se añade a ambos conjuntos el mismo atributo:

$$\text{Si } X \rightarrow Y \Rightarrow X \bullet Z \rightarrow Y \bullet Z$$

Dependencia transitiva: Si en la relación **T** existen las siguientes dependencias funcionales

$$\mathbf{X} \rightarrow \mathbf{Y} \ ; \ \mathbf{Y} \rightarrow \mathbf{Z} \ ; \mathbf{Y} \not\rightarrow \mathbf{X}$$

Se dice que **Z** *tiene una dependencia transitiva respecto* **X** *a través de* **Y**:

$$\mathbf{X} \rightarrow \mathbf{Z}$$

Unión o aditividad: Si **Y** depende de **X** y además **Z** también depende de **X** la composición de **Y** y **Z** depende también de **X**:

$$\text{Si} \quad \mathbf{X} \rightarrow \mathbf{Y} \ \text{ y } \ \mathbf{X} \rightarrow \mathbf{Z} \quad \Rightarrow \quad \mathbf{X} \rightarrow \mathbf{Y} \bullet \mathbf{Z}$$

Pseudo-transitividad: Si el atributo **Y** depende funcionalmente de **X** y el atributo **Z** depende funcionalmente del atributo compuesto **W•Y** se verifica entonces que **Z** depende funcionalmente del atributo compuesto **W•X**:

$$\text{Si} \quad \mathbf{X} \rightarrow \mathbf{Y} \ \text{ y } \ \mathbf{W} \bullet \mathbf{Y} \rightarrow \mathbf{Z} \quad \Rightarrow \quad \mathbf{W} \bullet \mathbf{X} \rightarrow \mathbf{Z}$$

Descomposición o proyectividad: Si **Y** depende funcionalmente de **X** y los valores de **Z** están incluidos en **Y** entonces **Z** depende funcionalmente de **X**:

$$\mathbf{X} \rightarrow \mathbf{Y} \ \text{ si } \quad \mathbf{Z} \subset \mathbf{Y} \ \Rightarrow \ \mathbf{X} \rightarrow \mathbf{Z}$$

2 ***Álgebra relacional*** .-

El modelo relacional es la forma de representar los datos y manipular dicha representación considerando la integridad de los datos. Estos, en el modelo relacional se representan mediante Relaciones y un procedimiento para su manipulación es el Álgebra Relacional.

El *Álgebra relacional* es un lenguaje de consulta procedimental.

Un *lenguaje de consulta* es aquel que el usuario utiliza para solicitar información de la base de datos. Este tipo de lenguajes puede ser *procedimental* o *no procedimental*. En el primer caso, el usuario indica la secuencia de operaciones que debe realizar sobre la base de datos para obtener el resultado deseado, mientras que en un lenguaje no procedimental el usuario describe la información que desea sin dar el procedimiento para determinarla.

Para realizar las consultas a la base de datos, el álgebra relacional dispone de un conjunto de operadores y una operación de asignación.

La *operación de asignación* asigna el valor de alguna expresión del álgebra a una relación nombrada.

Los *operadores* toman una o dos relaciones como entrada y producen una nueva operación como salida.

En álgebra relacional se definen ocho tipos de operadores básicos:

- *Operaciones de Conjuntos*: Unión, Intersección, Diferencia y Producto Cartesiano.
- *Operaciones relacionales*: Selección, Proyección, Reunión y División.

Además de estos operadores, se define el operador Renombrar que permite cambiar el nombre de los atributos de una relación.

3 ***Operador Renombrar*** **(ρ)**.-

Cambia el nombre de los atributos que sean necesarios en una relación antes de realizar una operación que pueda llevar a una relación con una cabecera en la que aparezcan dos atributos con el mismo nombre.

A partir de una relación especificada crea una nueva copia de ésta en la que sólo se han modificado los nombres de aquellos atributos que se quieren renombrar. La sintaxis, para una relación de nombre **R**, es:

$$\mathbf{R}\ \rho_{\text{Atributos Nuevos}}\ \textbf{(Atributos originales)}$$

4 ***Operaciones de conjuntos*.-**

Se dice que dos Relaciones son ***compatibles respecto a la unión*** sí y sólo sí sus cabeceras son idénticas. Esto implica que:

1.- Las dos tienen el mismo conjunto de nombres de atributos.

2.- Los atributos correspondientes se definen sobre el mismo dominio.

Se dice que dos Relaciones son ***compatibles respecto al producto*** sí y sólo sí sus cabeceras son disjuntas, esto es, no contienen nombres de atributos iguales.

Unión (∪): La unión de dos relaciones $\mathbf{R_1}$ y $\mathbf{R_2}$ compatibles respecto a la unión es una nueva relación **R** cuya cabecera es idéntica a la de las dos relaciones y cuyo cuerpo está formado por todas las tuplas pertenecientes a $\mathbf{R_1}$, a $\mathbf{R_2}$ o a las dos. Esto es, está formado por todas las tuplas que aparecen en cualquiera de las dos relaciones especificadas. Su sintaxis es:

$$\mathbf{R = R_1 \cup R_2}$$

Intersección ($\cap$): La intersección de dos relaciones $\mathbf{R_1}$ y $\mathbf{R_2}$ compatibles respecto a la unión es una nueva relación **R** cuya cabecera es idéntica a la de las dos relaciones y cuyo cuerpo está formado por todas las tuplas pertenecientes tanto a $\mathbf{R_1}$, como a $\mathbf{R_2}$ o a las dos. Esto es, está formado por todas las tuplas que aparecen en las dos relaciones especificadas. Su sintaxis es:

$$\mathbf{R = R_1 \cap R_2}$$

Diferencia (-): La diferencia de dos relaciones $\mathbf{R_1}$ y $\mathbf{R_2}$ compatibles respecto a la unión es una nueva relación **R** cuya cabecera es idéntica a la de las dos relaciones y cuyo cuerpo está formado por todas las tuplas pertenecientes a $\mathbf{R_1}$ pero no a $\mathbf{R_2}$. Esto es, está formado por todas las tuplas de la primera relación que no aparecen en la segunda. Su sintaxis es:

$$\mathbf{R = R_1 - R_2}$$

Producto Cartesiano (x): El producto cartesiano de dos relaciones $\mathbf{R_1}$ y $\mathbf{R_2}$ compatibles respecto al producto es una nueva relación **R** cuya cabecera es una combinación de las cabeceras de $\mathbf{R_1}$ y $\mathbf{R_2}$ y cuyo cuerpo está formado por el conjunto de todas las tuplas **t** tales que **t** es la combinación de la tupla $\mathbf{t_1}$ perteneciente a $\mathbf{R_1}$ y la tupla $\mathbf{t_2}$ perteneciente a $\mathbf{R_2}$. Esto es, está formado por todas las combinaciones posibles de tuplas, una de cada una de las dos relaciones. Su sintaxis es:

$$\mathbf{R = R_1 \times R_2}$$

Propiedades de los operadores de conjuntos

La Unión, la intersección y el producto cartesiano gozan de la propiedad asociativa y de la conmutativa, pero no así la diferencia. Así si $\mathbf{R_1}$, $\mathbf{R_2}$ y $\mathbf{R_3}$ son relaciones arbitrarias (que cumplen con la compatibilidad respecto de la operación a realizar) se tendrá:

Propiedad Asociativa:

Unión .- $\mathbf{(R_1 \cup R_2) \cup R_3 \Leftrightarrow R_1 \cup (R_2 \cup R_3) \Leftrightarrow R_1 \cup R_2 \cup R_3}$

Intersección.- $\mathbf{(R_1 \cap R_2) \cap R_3 \Leftrightarrow R_1 \cap (R_2 \cap R_3) \Leftrightarrow R_1 \cap R_2 \cap R_3}$

Producto Cartesiano .- $\mathbf{(R_1 \times R_2) \times R_3 \Leftrightarrow R_1 \times (R_2 \times R_3) \Leftrightarrow R_1 \times R_2 \times R_3}$

Propiedad Conmutativa:

Unión .- $\mathbf{R_1 \cup R_2 \Leftrightarrow R_2 \cup R_1}$

Intersección.- $\mathbf{R_1 \cap R_2 \Leftrightarrow R_2 \cap R_1}$

Producto Cartesiano .- $\mathbf{R_1 \times R_2 \Leftrightarrow R_2 \times R_1}$

5 ***Operaciones Relacionales*** **.-**

Selección (σ): La selección de tuplas de una relación **R** es otra relación con la misma cabecera que **R** y cuyo cuerpo está formado por las tuplas de **R** que verifican una ***condición*** entre atributos.

En la condición pueden aparecer operadores de comparación (=, <>, >=, etc.) y booleanos (and, or, not). Los atributos que aparecen en la condición deben estar definidos sobre el mismo dominio. La sintaxis del operador es:

$$\sigma_{Condición} (R)$$

Proyección (Π): La proyección de la Relación R según los atributos $A_1, A_2, \ldots, A_n$ es otra relación que tiene por cabecera los atributos indicados y en cuyo cuerpo aparecen todas las tuplas de R restringidas a dichos atributos, eliminando tuplas repetidas. La sintaxis del operador es:

$$\Pi_{A1,\ a2,\ \ldots An} (R)$$

Reunión: La reunión de dos relaciones específicas es otra relación que contiene todas las posibles combinaciones de tuplas, una de cada una de las dos relaciones, tales que las dos tuplas participantes de la combinación satisfacen una condición especificada. Se consideran dos tipos de reunión:

Reunión natural (*): Sea ($A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_m$) la cabecera de la relación R_1 y ($B_1, B_2, \ldots, B_m, C_1, C_2, \ldots, C_p$) la cabecera de la Relación R_2 estando los atributos del mismo nombre definidos en el mismo dominio.

Si se consideran los tres atributos compuestos **A, B, C**, la reunión natural de R_1 y R_2 es una relación con la cabecera (**A, B, C**) y un cuerpo formado por todas las tuplas **t** (**A:a, B:b, C:c**) tales que **t** es la combinación de la tupla t_1 que aparece en R_1 con el valor **a** en **A** y el valor **b** en **B**, y la tupla t_2 que aparece en R_2 con el valor **b** en **B** y el valor **c** en **C**. La sintaxis del operador es:

$$R_1 * R_2$$

Si las dos relaciones no tienen atributos en común, la reunión natural coincide con el producto cartesiano:

$$R_1 * R_2 \Leftrightarrow R_1 \times R_2$$

Goza de las propiedades asociativa y conmutativa:

Asociativa: $(R_1 * R_2) * R_3 \Leftrightarrow R_1 * (R_2 * R_3) \Leftrightarrow R_1 * R_2 * R_3$

Conmutativa: $R_1 * R_2 \Leftrightarrow R_2 * R_1$

Reunión Theta (|x|): Permite reunir dos relaciones en función de una condición diferente de la igualdad. Sean las relaciones R_1 y R_2 compatibles respecto al producto y sea θ un operador. La reunión theta de la relación R_1 según el atributo **A** con la relación R_2 según el atributo **B** es una relación con la misma cabecera que el producto cartesiano de R_1 y R_2 y un cuerpo formado por el conjunto de todas las tuplas **t** tales que **t** pertenece al producto cartesiano si la evaluación de la condición **A θ B** resulta verdadera. Los atributos **A** y **B** deben estar definidos sobre el mismo dominio y la operación θ debe ser aplicable a ese dominio. La sintaxis del operador es:

$$\sigma_{A\,\theta\,B}\,(R_1 \times R_2) \Leftrightarrow R_1\ |x|_{A\,\theta\,B}\ R_2$$

División (÷): Sea (A_1, A_2,..., A_n, B_1, B_2,..., B_m) la cabecera de la relación R_1 y (B_1, B_2,..., B_m) la cabecera de la Relación R_2 estando los atributos del mismo nombre definidos en el mismo dominio.

Si se consideran los atributos compuestos **A** y **B**, la división de R_1 (dividendo) entre R_2 (divisor), es otra relación con la cabecera (**A**) y un cuerpo formado por el conjunto de todos los valores de R_1 en el atributo **A**, cuyos valores correspondientes en el atributo **B** incluyen a todos los valores del atributo **B** en la relación R_2.

La división toma dos relaciones, una binaria y otra unaria, y construye una relación formada por todos los valores de un atributo de la relación binaria que concuerdan , en el otro atributo, con todos los valores en la relación unaria. La sintaxis del operador es:

$$R_1 \div R_2$$

6 ***Operaciones adicionales*** .-

Ampliación (α): Toma una relación **R** y crea otra nueva relación con un atributo mas que la original cuyos valores se obtienen evaluando alguna expresión de cálculo escalar. La sintaxis del operador es:

$$R\ \alpha_{\text{Cálculo Escalar}}\ \text{(nombre atributo)}$$

Resumen (Ω): Permite incorporar operaciones de agregados (cuenta, suma, promedio, máximo, mínimo, etc.). A partir de una relación **R** y de una lista de atributos, obtiene otra relación en cuya cabecera aparecen los atributos de R especificados y un nuevo atributo, con el nombre indicado, siendo los valores de éste último el resultado de evaluar la expresión de agregados. La sintaxis del operador es:

$$R\ \text{(lista atributos)}\ \Omega_{\text{Cálculo Agregados}}\ \text{(nombre atributo)}$$

División generalizada (÷): Dadas la relación $\mathbf{R_1}$ con la cabecera (**A, B**) y la relación $\mathbf{R_2}$ con la cabecera (**B, C**) donde los atributos **A, B, C** pueden ser compuestos, la división generalizada produce una relación que tiene como cabecera (**A, C**) y un cuerpo formado por todas las tuplas (**A:a, C:c**) tales que aparece una tupla (**A:a, B:b**) en $\mathbf{R_1}$ para todas las tuplas (**B:b, C:c**) que aparecen en $\mathbf{R_2}$.

Si **C** está vacío la operación se reduce a la *división* entre $\mathbf{R_1}$ y $\mathbf{R_2}$

Si **A** está vacío la operación se reduce a la *división* entre $\mathbf{R_2}$ y $\mathbf{R_1}$

Si **B** está vacío la operación degenera en el *producto cartesiano* de $\mathbf{R_1}$ y $\mathbf{R_2}$

Al ser la división un caso particular de la división generalizada utilizan la misma simbología. La sintaxis del operador es:

$$\mathbf{R_1 \div R_2}$$

7 ***Asignación relacional*** .-

El objetivo de esta operación es asignar, mediante una ***etiqueta***, el valor de alguna expresión algebraica. La asignación se utiliza en consultas que requieren una expresión algebraica extensa.

Etiqueta ← Expresión