
Cálculo Relacional

Fundamentos de Bases de Datos

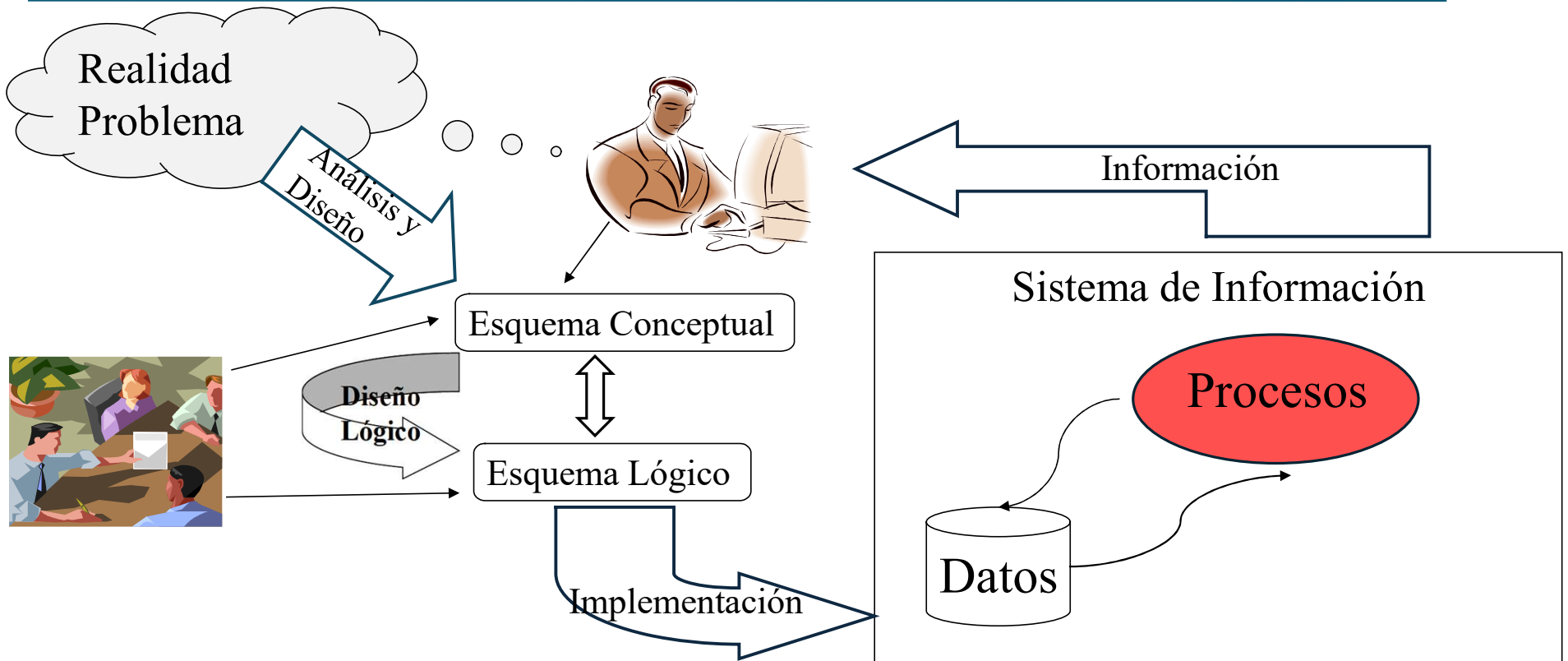
Inco – Fing - Udelar

Cálculo Relacional

- **Temas**

- Lenguajes de Consulta
- Cálculo Relacional de Tuplas
- Fórmulas Seguras
- Cálculo Relacional de Dominios

Construcción de un Sistema de Información



Lenguajes de Consulta en MR

- **Lenguajes de Consulta (QL) en Modelo Relacional**
- **La idea básica de los lenguajes de consulta es tomar las tablas de la base y construir nuevas tablas “al vuelo”.**
- **Para cada constructor del lenguaje, típicamente, se define:**
 - Cuál es el esquema del resultado.
 - Cuál es el conjunto de tuplas del resultado.

Lenguajes de Consulta en MR

- **Para definir un lenguaje, se deben considerar dos aspectos:**
 - Sintaxis:
 - Una descripción de cuáles son las frases correctamente escritas en ese lenguaje.
 - Semántica:
 - Una forma de establecer la correspondencia de cada frase bien escrita según la sintaxis con un significado dado.
- **Hay otro aspecto que va a cobrar importancia en los QL:**
 - Pragmática:
 - Cómo el contexto influye en el uso del lenguaje. El contexto está dado por el esquema (Estructuras y Restricciones) y el significado de ese esquema para quienes interpretan los datos.

Cálculo Relacional - Introducción

- **De qué se trata ?**

- Familia de lenguajes de consulta sobre Modelo Relacional
- Basado en fórmulas de Lógica de 1er orden para definir conjuntos
- Una consulta en CR es una especificación por comprensión de un conjunto de tuplas.

- **Dos Sublenguajes:**

- Cálculo de Tuplas: variables de tipo tupla, es decir, el universo está formado por tuplas.
- Cálculo de Dominios: variables por valor de atributo, es decir, el universo está formado por valores individuales.
- Sólo se va a trabajar con el primero de estos lenguajes.

Cálculo de Tuplas - Sintaxis

- **Las expresiones del cálculo son de la forma:**
 - $\{ \langle t_1, \dots, t_n \rangle / \varphi(x_1, \dots, x_n) \}$ donde:
 - t_i son términos de la forma $x_i.A_k$ o bien de la forma c_i . Aquí, x_i es una variable libre de φ y A_k un atributo de una tabla para la que x_i representa una tupla y c_i es una constante de algún dominio (número, string, fecha, etc.)
 - φ es una fórmula donde las x_i que aparecen a la izquierda de la barra son exactamente las variables libres de φ . ($FV(\varphi) = \{x_1, \dots, x_m\}$).
- **Hay que definir el lenguaje de las fórmulas φ válidas.**
 - Es un lenguaje de primer orden, en donde los predicados son las tablas de la base, consideradas como unarias.
 - Para definirlo, hay que definir dos lenguajes: Los términos y las fórmulas propiamente dichas.

Cálculo de Tuplas - Sintaxis

- **Términos:**

- x_i es un término, si x_i es una variable.
- c_i es un término, si c_i es una constante de algún dominio.
- $x_i.A_j$ es un término, si x_i es una variable y A_j es un nombre de atributo.
- Estos son todos los términos. (no hay más funciones que la selección de atributos).

- **Hay dos tipos de términos:**

- Atómicos: c_i o $x_i.A_j$
- No atómicos: x_i (Variables)

Cálculo de Tuplas - Sintaxis

- **Las fórmulas pueden ser:**

- $t_i <op> t_j$ donde t_i y t_j son términos atómicos ($x_i.A_k$ o c_j) y $<op> \in \{ =, <, >, \neq, \dots \}$.
- $P_i(x_i)$ es una fórmula si P_i es el nombre de una tabla y x_i es una variable.
- $(\varphi <BinConn> \psi)$ es una fórmula si φ y ψ son fórmulas y $<BinConn> \in \{ \wedge, \vee, \rightarrow, \leftrightarrow \}$.
- $(\neg \varphi)$ es una fórmula si φ es una fórmula.
- $(\exists x_i \varphi)$ y $(\forall x_i \varphi)$ son fórmulas si φ es una fórmula y x_i es una variable.
- Estas son todas las fórmulas.

Cálculo de Tuplas - Ejemplos

- **Ejemplo 1.**

- Dado el siguiente Esquema Relacional:

FABS (numF, Nombre, Direcc)

PRODS (numP, desc)

VENTAS(numF, numP, precio)

- ¿Qué ventas fueron por un precio mayor de \$50.000?
- Expresión en CRT:

$\{t / \text{VENTAS}(t) \wedge t.\text{precio} > 50000\}$

- donde: VENTAS(t) especifica que t pertenece a VENTAS y
t.precio referencia al atributo PRECIO de la variable de tupla t.

- Resultado:

- tuplas t de VENTAS con t.precio > 50000

Cálculo de Tuplas - Ejemplos

- **Ejemplo 2.**

$$\{ \langle t.\text{Nombre}, t.\text{Direc} \rangle / \text{FABS}(t) \wedge \\ \exists v (\text{Ventas}(v) \wedge v.\text{numF} = t.\text{numF} \wedge \\ \neg \exists v1 (\text{Ventas}(v1) \wedge v1.\text{precio} > v.\text{precio})) \\ \}$$

- Nombre y dirección de los Fabricantes que venden un producto con el precio máximo.

Cálculo de Tuplas - Semántica

- **El universo está formado por todas las tuplas que se pueden construir con todos los dominios.**
- **El resultado de una consulta está dado por todas las tuplas del universo que cumplen con la condición expresada en la fórmula de la consulta.**

Cálculo de Tuplas – Semántica

- **La expresión $\{ \langle x_1.A_k, \dots, x_n.A_j \rangle / \varphi \}$ define una relación tal que:**
 - El esquema tiene los n atributos A_k, \dots, A_j
 - La instancia es la formada por todas las tuplas t del universo que tienen aridad n , tales que $\varphi(t)$ es verdadera.
- **Observaciones:**
 - La tabla del resultado tiene exactamente n atributos con los nombres A_k, \dots, A_j exactamente en ese orden.

Cálculo de Tuplas – Semántica

- **Las fórmulas son verdaderas si cumplen:**
 - φ de la forma $P(x_i)$ es verdadera, si y solo si hay alguna tupla t en el universo tal que t pertenece a la tabla P en la base.
 - φ de la forma $t_i \langle op \rangle t_j$ es verdadera, si y solo si se cumple la relación especificada por $\langle op \rangle$ entre los valores asociados a t_i y t_j en el momento de la consulta. (Recordar que t_i y t_j son términos atómicos: representan valores de atributos o constantes de un dominio.)
 - El resto de las fórmulas se evalúan según la semántica clásica de la lógica de primer orden.

Cálculo de Tuplas – Universo.

- **Observar que el universo siempre es el conjunto de todas las tuplas que es posible formar con todos los dominios de los atributos.**
 - Es un conjunto infinito.
 - Contiene a tuplas de cualquier aridad.
 - Contiene a tuplas con cualquier dominio en cualquier posición.
 - Se deben tener en cuenta estas características a la hora de escribir las consultas.

Cálculo de Tuplas - Ejemplos

- **Fabs(numF, Nombre, Direcc), Prods(numP, Desc), Ventas(numF, numP, Precio).**
- **Consulta 1:**
 - Dar nombre y dirección de fabricantes que hicieron ventas por más de \$1000.

$$\{ \langle t.\text{nombre}, t.\text{Direcc} \rangle / \text{FABS}(t) \wedge \exists v (\text{VENTAS}(v) \wedge v.\text{numF} = t.\text{numF} \wedge v.\text{precio} > 1000) \}$$

- **Observaciones:**
 - Se asumen las reglas de eliminación de paréntesis de la lógica de primer orden.
 - Las variables libres en una expresión del cálculo relacional deberán ser las que aparezcan a la izquierda de la barra (/).
 - En la consulta, t es la única variable libre, y v es ligada.

Cálculo de Tuplas - Ejemplos

- **Consulta 2:**

Fabs(numF, Nombre, Direcc), Prods(numP, Desc), Ventas(numF, numP, Precio)

- Dar parejas de numero de fabricante que vendieron los mismos productos.

$$\{ \langle v1.numF, v2.numF \rangle / VENTAS(v1) \wedge VENTAS(v2) \wedge v1.numP = v2.numP \wedge v1.numF > v2.numF \}$$

- **Observaciones:**

- Dos o más variables de tupla de una consulta pueden abarcar la misma relación. En este caso v1 y v2. Esto permite evitar el uso de posiciones.

Cálculo de Tuplas - Ejemplos

- **Consulta 3:**

- Dar los números de fabricantes, la descripción de los productos que venden y a qué precio lo hacen.

$\{ \langle t1.numF, t2.desc, t1.precio \rangle / VENTAS(t1) \wedge PRODS(t2) \wedge t1.numP = t2.numP \}$

- **Consulta 4:**

- Dar los nombres de los fabricantes que vendieron productos con descripción d1.

$\{ f.nombre / FABS(f) \wedge \exists p (PRODS(p) \wedge p.desc='d1' \wedge \exists v (VENTAS(v) \wedge f.numF = v.numF \wedge p.numP = v.numP)) \}$

Cálculo de Tuplas - Ejemplos

- **Consulta 5:**

- Dar los nombres de los fabricantes que venden todos los productos.

~~$\{ f.\text{nombre} / \text{FABS}(f) \wedge \forall p (\text{PRODS}(p) \wedge \exists v (\text{VENTAS}(v) \wedge f.\text{numF} = v.\text{numF} \wedge p.\text{numP} = v.\text{numP})) \}$~~

- Observar que:
 - p es cualquier tupla del universo.
 - El universo es infinito.
 - La consulta no es correcta.

Cálculo de Tuplas - Ejemplos

- **Consulta 5:**

- Dar los nombres de los fabricantes que venden todos los productos.

- El lenguaje NO PERMITE escribir algo así:

$\{ f.\text{nombre} / f \in \text{FABS} \wedge \forall p \in \text{PRODS.}$

$\exists v \in \text{VENTAS.}$

$(f.\text{numF} = v.\text{numF} \wedge p.\text{numP} = v.\text{numP}) \}$

- Sin embargo, se pueden escribir fórmulas equivalentes mediante las siguientes traducciones (Relativización):

- $\exists t \in P. \varphi \equiv \exists t (P(t) \wedge \varphi)$

- $\forall t \in P. \varphi \equiv \forall t (P(t) \rightarrow \varphi)$

Cálculo de Tuplas - Ejemplos

- **Consulta 5:**

- Dar los nombres de los fabricantes que venden todos los productos.

$$\{ \text{f.nombre} / \text{FABS}(\text{f}) \wedge \\ \forall \text{p} (\text{PRODS}(\text{p}) \rightarrow \exists \text{v} (\text{VENTAS}(\text{v}) \wedge \text{f.numF} = \text{v.numF} \wedge \text{p.numP} = \text{v.numP}) \\) \\ \}$$

- **Observación:**

- Pensamos relativizado (con \in), escribimos con la forma sin relativizar (se escribe con la traducción).

Cálculo de Tuplas – Ejemplos

- **Consulta 6:**

- Dar los nombres de fabricantes que vendieron todos los productos con descripción 'd3'.

~~$\{ t.Nombre / FABS(t) \wedge ((\forall p)(PRODS(p) \wedge p.desc = 'd3' \rightarrow (\exists v)(VENTAS(v) \wedge p.numP = v.numP \wedge t.numF = v.numF))) \}$~~

- Qué resultado debe dar la consulta si no hay productos con descripción 'd3' ?
 - En este caso debería dar vacío, por lo que la consulta anterior es incorrecta.

$\{ t.Nombre / FABS(t) \wedge ((\forall p)(PRODS(p) \wedge p.desc='d3' \rightarrow (\exists v)(VENTAS(v) \wedge p.numP = v.numP \wedge t.numF = v.numF))) \wedge (\exists q)(PRODS(q) \wedge q.desc='d3') \}$

Cálculo de Tuplas – Fórmulas Inseguras

- **Fórmulas inseguras**

- Fórmulas que permiten resultados infinitos en consultas CRT.
- Cómo puede pasar esto ?
 - En la medida de que el universo de valores es infinito, pueden construirse consultas CRT que generen resultados infinitos.
 - Ejemplos:
 - $\{ t / \neg \text{FABS}(t) \}$ Es insegura ya que retorna las tuplas del universo infinito que no pertenecen al conjunto finito de Fabricantes.
 - $\{ t / 1=1 \}$ Esta consulta no respeta la sintaxis presentada ya que la variable t no está libre en la expresión. Si fuera posible escribirla, entonces sería insegura dado que t tomaría como valores todo el universo.

Cálculo de Tuplas - Fórmulas Seguras

- **Dominio de una Expresión CRT:**

Es el conjunto de todos los valores que, aparecen como valores constantes en la expresión o bien existen en cualquier tupla de las relaciones a las que se hace referencia en la expresión.

- **Fórmulas seguras:**

Se dice que una expresión CRT es segura si todos los valores de su resultado pertenecen al dominio de la expresión.

Cálculo de Tuplas - Fórmulas Seguras

- **Ejemplos**

$\{v.numF / \neg VENTAS(v)\}$

- Es insegura porque el resultado incluye valores que no están en el Dominio de la expresión (formado por las tuplas de VENTAS).

$\{t.numF / FABS(t) \wedge \neg(\exists v)(VENTAS(v) \text{ and } v.numF = t.numF)\}$

- Es segura, porque el resultado solo incluye valores de FABS.

$\{t / (\exists f) (FABS(f))\}$

- Es insegura porque el resultado incluye valores que no están en el Dominio de la expresión (formado por las tuplas de FABS).

Cálculo de Tuplas - Fórmulas Seguras

- **Algunos criterios para chequear si la fórmula es segura**

- Pensar en la fórmula de la consulta, traducida con \wedge, \vee y \neg .
- Si la fórmula tiene la forma $\varphi_1 \vee \dots \vee \varphi_n$, en cada φ_i deben aparecer todas las variables libres en un predicado no negado.
- Si la fórmula es de la forma $\varphi_1 \wedge \dots \wedge \varphi_n$, cada variable debe aparecer en al menos una φ_i en un predicado no negado.
- Recordar las asociaciones de \exists con \vee y \forall con \wedge y razonar de forma similar.

- Ejs:
 - $\{ f / \neg \text{FABS}(f) \}$
 - $\{ f / \exists p (\text{FABS}(f) \rightarrow \text{PRODS}(p) \wedge p.\text{numF} = f.\text{numF}) \}$
 - $\{ t.\text{numF} / \neg \text{FABS}(t) \vee \text{VENTAS}(t) \}$

- Son todas expresiones a evitar.

Cálculo de Dominios

- **El Cálculo Relacional de Dominios:**

- Es otro lenguaje para BDs relacionales basado en el cálculo de predicados.
- El universo de las interpretaciones es la union de todos los dominios. (Diferente que en CRT).
- Las fórmulas son del tipo: $\{t_1, \dots, t_n / \varphi\}$
 - donde:
 - t_i es o una variable o una constante. Si es una variable, entonces aparece libre en φ .
 - Todas las variables libres de φ aparecen a la izquierda.
 - φ es una fórmula de primer orden.

Cálculo de Dominios

- **Usando el CRD:**

- Permite expresiones más compactas que las de CRT, debido a que se pueden expresar igualdades a través del matcheo de variables de atributos.

- Ejemplo:

$\{ \text{nom, dir} / (\exists \text{nf}) (\text{FABS}(\text{nf, nom, dir}) \wedge \text{VENTAS}(\text{nf, _, _})) \}$

- Observar que:

- Las igualdades no son necesarias (aunque se pueden usar).
- Los atributos que no se necesitan se pueden poner con “_”.
- Igualmente se podrían poner constantes en atributos.

Cálculo de Dominios

- **Ejemplos:**

- Dar descripciones de productos vendidos a \$100.

$\{ \text{Descrip} / (\exists np) (\text{PRODS}(np, \text{Descrip}) \wedge \text{VENTAS}(_, np, 100)) \}$

- Dar descripciones de productos vendidos sólo por Pepe.

$\{ \text{Descrip} / (\exists np) (\text{PRODS}(np, \text{Descrip}) \wedge$
 $(\forall nf) (\text{VENTAS}(nf, np, _) \rightarrow (\text{FABS}(nf, \text{”Pepe”}, _)) \wedge \text{VENTAS}(_, np, _)) \}$

- Dar nombre de fabricantes que nunca hicieron ventas por mas de \$100.

$\{ \text{Nom} / (\exists nf) (\text{FABS}(nf, \text{Nom}, _) \wedge \neg (\exists p, np) (\text{VENTAS}(nf, np, p) \wedge p > 100)) \}$

Consultas en CR: Pragmática

Docentes(CI, nombre, dir, tel)

Estudiantes(CI , nombre, dir, tel)

- **Dar el nombre y la dirección de los estudiantes que son docentes.**

~~{ <e.nombre, e.dir> / Estudiantes(e) ∧
∃d(Docente(d) ∧ e.nombre = d.nombre) }~~

{ <e.nombre, e.dir> / Estudiantes(e) ∧
∃d(Docente(d) ∧ e.CI = d.CI) } ✓

Consultas en CR: Pragmática

- **Al realizar las consultas, se accede a las claves de un elemento y luego se relacionan elementos a través de esas claves.**
 - Pueden existir excepciones, pero son muy poco habituales.

Equivalencias lógicas útiles

- **Entre cuantificadores:**

- $(\forall x) (P(x)) \leftrightarrow \neg (\exists x) (\neg (P(x)))$
- $(\exists x) (P(x)) \leftrightarrow \neg (\forall x) (\neg (P(x)))$

- **Entre conectivos:**

- $\neg(x \vee y) \leftrightarrow (\neg x) \wedge (\neg y)$
- $\neg(x \wedge y) \leftrightarrow (\neg x) \vee (\neg y)$
- $x \rightarrow y \leftrightarrow \neg x \vee y$