



Dpto. Ciencias e Ingeniería de la Computación
 Universidad Nacional del Sur

ELEMENTOS DE BASES DE DATOS

Segundo Cuatrimestre 2013

Clase 5: Modelo Relacional – Lenguaje de Consulta Relacional – Calculo Relacional de Tuplas (CRT)



Mg. María Mercedes Vitturini
 [mvitturini@uns.edu.ar]

Lenguajes de Consulta

Servicio DBMS: lenguaje de alto nivel para gestionar datos

- **Lenguaje de Definición de Datos (LDD/DDL):** considera las sentencias para especificar *el esquema de una base de datos*.
- **Lenguaje de Manipulación de Datos (LMD/DML):** considera las sentencias para *modificar y recuperar información almacenada* en una base de datos, ie. manipular instancias.
 - Categorías de lenguajes DML puros:
 - **Algebraicos o procedurales:** las consultas son expresadas aplicando operadores específicos a las relaciones.
 - **Lógicos o declarativos:** denominadas **cálculo relacional**, expresadas por fórmulas lógicas.

EBD2012_6 - Mg. Mercedes Vitturini

Cálculo Relacional de Tuplas

- Las **consultas en CRT se expresan en forma declarativa**, sin especificar el procedimiento efectivo para obtener la información.
- Formato general de una consulta:

$$\{ t : P(t) \}$$

Leer : "El conjunto de tuplas *t* tal que satisfacen el predicado *P*".
- El **resultado es una relación** integrada por el conjunto de tuplas *t* que satisfacen el predicado *P*.



EBD2012_6 - Mg. Mercedes Vitturini

nroCarrera	nroCaballo	posición
CC-01	CB-01	1
CC-01	CB-02	null
CC-01	CB-03	2
CC-01	CB-04	1
CC-01	CB-05	3
CC-02	CB-01	Null
CC-02	CB-03	1
CC-02	CB-04	3
CC-02	CB-05	2
CC-03	CB-01	2
CC-03	CB-02	1
CC-03	CB-03	Null
CC-03	CB-04	3
CC-03	CB-05	4
CC-04	CB-01	3
CC-04	CB-02	4
CC-04	CB-04	null
CC-04	CB-05	2

corre

Ejemplo

nroCarrera	nombre	fecha
CC-01	Premier	11-04-2012
CC-02	Trote	31-06-2012
CC03	Speed	01-08-2012
CC-04	Millón	27-08-2012

carrera

nroCaballo	nombre	edad	peso
CB-01	Indio	9	280
CB-02	Orso	6	289
CB-03	Juana	2	230
CB-04	Grecia	8	245
CB-05	Chino	8	256
CB-06	Cristal	5	180
CB-07	Kaiser	3	205

caballo

EBD2012_6 - Mg. Mercedes Vitturini

Participa una relación:

CRT – Ejemplos

"Caballos de más de ocho años?"

$$\{ t \mid t \in \text{caballo} \wedge t[\text{edad}] > 8 \}$$

Selección

"Nombre de las carreras que se corrieron en agosto de 2012?"

$$\{ t \mid \exists c \in \text{carrera} (t[\text{nombre}] = c[\text{nombre}] \wedge c[\text{fecha}] >= "01/08/2012" \wedge c[\text{fecha}] <= "31/08/2012") \}$$

Selección + Proyección

Participa más relaciones:

Nombre de fantasía de caballos que corrieron alguna carrera?

$$\{ t \mid \exists cb \in \text{caballo}, \exists co \in \text{corre}, (t[\text{nombre}] = cb[\text{nombre}] \wedge cb[\text{nroCaballo}] = co[\text{nroCaballo}]) \}$$

Join Natural + Selección + Proyección

EBD2012_6 - Mg. Mercedes Vitturini

CRT – Notación

- $t[A]$ denota el valor de la tupla *t* en el atributo *A*.
- $t \in r$ denota que la **tupla *t*** está en **relación de pertenencia** con la **relación *r***.
- \forall : cuantificador universal "para todo".
- \exists : cuantificador existencial "existe".
- \wedge : conjunción lógica "y".
- \vee : disyunción lógica "o".
- \neg : negación lógica "no".
- \Rightarrow : implicación o entonces.

EBD2012_6 - Mg. Mercedes Vitturini

Definición Formal del CRT

- Dada una expresión en el CRT de la forma:
 $\{t : P(t)\}$
- P es una **fórmula** que puede contener *variables de tupla*.
- Una variable de tupla puede estar *libre* o *ligada*.
- Una variable está **ligada** cuando está cuantificada universalmente (\forall) o existencialmente (\exists).
- **Ejemplo:**
 $\{t : \exists s \in \text{cliente } (s[\text{localidad}] = \text{"Bahia Blanca"} \wedge t=s)\}$
 – En el ejemplo t es una variable **libre** y s está **ligada**.

EBD2012_6 - Mg. Mercedes Vitturini

Definición Formal del CRT

- Una **fórmula** del CRT *está compuesta de átomos*.
- Un **átomo** puede ser de las siguientes formas:
 - $t \in r$ donde t es una variable tupla y r una relación.
 - $t[X] \theta u [Y]$ donde:
 - t y u son variables de tupla,
 - X atributos definidos en t , Y atributos definidos en u ,
 - θ un operador relacional del conjunto $\{<, \leq, >, \geq, =, \neq\}$.
 - $t[X] \theta c$ donde:
 - t es una variable tupla, θ un operador relacional y c una constante en el dominio del atributo X .

EBD2012_6 - Mg. Mercedes Vitturini

Definición Formal del CRT

Las **fórmulas** se *construyen usando las reglas*:

- Un átomo es una fórmula.
- Si P es una fórmula entonces lo son $\neg P$ y (P) .
- Si P_1 y P_2 son fórmulas entonces lo son $P_1 \wedge P_2$, $P_1 \vee P_2$ y $P_1 \Rightarrow P_2$.
- Si $P(s)$ es una fórmula y r es una relación, entonces son fórmulas:
 - $\exists s \in r (P(s))$ y
 - $\forall s \in r (P(s))$.

EBD2012_6 - Mg. Mercedes Vitturini

Significado de expresiones

- $\exists t \in r (Q(t))$ significa *"existe una tupla t en la relación r tal que el predicado $Q(t)$ es verdadero"*.
- $\forall t \in r (Q(t))$ significa *"el predicado $Q(t)$ es verdadero para todas las tuplas t en la relación r "*.
- $P_1 \wedge P_2$ es equivalente a $\neg(\neg P_1 \vee \neg P_2)$.
- $P_1 \Rightarrow P_2$ es equivalente a $\neg P_1 \vee P_2$.
- $\forall t \in r (Q(t))$ es equivalente a $\neg(\exists t \in r (\neg Q(t)))$.

EBD2012_6 - Mg. Mercedes Vitturini

Consultas en CRT

- Nombre de clientes que tienen un préstamo en la sucursal Universitario:

$$\{t : \exists s \in \text{presta}, \exists u \in \text{préstamos} \\ (u [\text{Nro-Prest}] = s [\text{Nro-Prest}] \\ \wedge u [\text{Nom-Sucur}] = \text{"Universitario"} \\ \wedge t [\text{Nom-Cliente}] = s [\text{Nom-Cliente}]) \}$$
 - Se lee como "el conjunto de las tuplas (con atributo **Nom-Cliente**) para las cuales el **Cliente** tiene un préstamo en la sucursal Universitario".
 - La variable tupla u asegura que el cliente tiene un préstamo en la sucursal Universitario.

EBD2012_6 - Mg. Mercedes Vitturini

Consultas en CRT

- Clientes que tienen un préstamo, una cuenta, o ambos en un banco:

$$\{t : \exists s \in \text{presta } (t[\text{Nom-Cliente}] = s[\text{Nom-Cliente}]) \vee \\ \exists u \in \text{deposita } (t[\text{Nom-Cliente}] = u[\text{Nom-Cliente}]) \}$$
 - Esto es, el conjunto de tuplas con atributo **Nom-Cliente** para las cuales vale una de las siguientes condiciones:
 - **Nom-Cliente** aparece en alguna tupla de la relación **presta** como prestatario del banco.
 - **Nom-Cliente** aparece en alguna tupla de la relación **deposita** como depositante del banco.

EBD2012_6 - Mg. Mercedes Vitturini

Consultas en CRT

- Clientes que tienen un préstamo y una cuenta en el banco:
 $\{ t : \exists s \in \text{presta}(t[\text{Nom_Cliente}]=s[\text{Nom_Cliente}]) \wedge \exists u \in \text{deposita}(t[\text{Nom_Cliente}]=u[\text{Nom_Cliente}]) \}$
- Clientes que tienen un préstamo **pero no** una cuenta en el banco:
 $\{ t : \exists s \in \text{presta}(t[\text{Nom_Cliente}]=s[\text{Nom_Cliente}]) \wedge \neg(\exists u \in \text{deposita}(s[\text{Nom_Cliente}]=u[\text{Nom_Cliente}])) \}$

EBD2012_6 - Mg. Mercedes Vitturini

El cuantificador \forall

$$\forall t \in r (Q(t))$$

- Significa que $Q(t)$ es verdadero para todas las tuplas t en r

$$P \Rightarrow Q$$

- “P implica Q”, esto es, si P es verdadero, entonces Q debe ser verdadero, ó $\neg P \vee Q$.

Ejemplo

- Nombres de Cliente con cuenta en **todas** las sucursales de Tres Arroyos
 – Si existe algún clientes que “para todas las sucursales, si la sucursal es de Tres Arroyos, entonces el cliente tiene una cuenta en dicha sucursal”

EBD2012_6 - Mg. Mercedes Vitturini

Consultas en CRT

- Clientes que tienen cuenta en **todas** las sucursales situadas en Tres Arroyos:
 $\{ t : \exists cl \in \text{clientes} (t[\text{Nom-Cliente}]=cl[\text{Nom-Cliente}] \wedge (\forall suc \in \text{sucursales} (suc[\text{Ciu-Sucur}] = \text{"Tres Arroyos"} \Rightarrow \exists dep \in \text{deposita}, \exists cta \in \text{cuentas} (dep[\text{Nom-Cliente}]=cl[\text{Nom-Cliente}] \wedge (dep[\text{Nro-Cuenta}]=cta[\text{Nro-Cuenta}] \wedge cta[\text{Nom-Sucur}] = suc[\text{Nom-Sucur}]))))) \}$

EBD2012_6 - Mg. Mercedes Vitturini

Seguridad de las Expresiones

- Supongamos que deseamos escribir una consulta del tipo:
 $\{ t : \neg(t \in \text{préstamo}) \}$
- Existen infinitas tuplas que no están actualmente en **préstamo**, por lo que no se desean estos tipos de expresiones.
- El **dominio** de una fórmula P en CRT es el conjunto de todos los valores que aparecen **explícitamente** en P.
 – $\text{dom}(t \in \text{préstamos} \wedge t[\text{Monto}] > 1200)$ es el conjunto de valores que aparecen en préstamos.
 – $\text{dom}(\neg(t \in \text{préstamos}))$ es el conjunto de valores que no están en préstamos.

EBD2012_6 - Mg. Mercedes Vitturini

Seguridad de las Expresiones

- Una expresión $\{ t : P(t) \}$ es **segura** si todos los valores que aparecen en el **resultado** son valores de $\text{dom}(P)$.
- La expresión $\{ t : \neg(t \in \text{préstamos}) \}$ no es segura.
 – Sabemos que $\text{dom}(\neg(t \in \text{préstamos}))$ es el conjunto de valores que aparecen en préstamos.
 – Sin embargo, es posible tener una tupla t que no esté en préstamo que contenga valores que no aparecen en préstamo.
- El resto de las expresiones que presentamos hasta ahora eran seguras.

EBD2012_6 - Mg. Mercedes Vitturini

LMD - Resumen

- Dos notaciones diferentes o categorías de LMD
 - **Algebraicas o procedurales**: las consultas son expresadas aplicando operadores específicos a las relaciones.
 - Álgebra Relacional (AR)
 - **Lógicas o declarativas**: denominadas **cálculo relacional**, expresadas por fórmulas lógicas.
 - Cálculo Relacional de Tuplas
 - Cálculo Relacional de Dominios
- Lenguajes Comerciales
 - **Structured Query Language (SQL)**

Técnicos Puros

Toma características de ambos

EBD2012_6 - Mg. Mercedes Vitturini

Ejercicios propuestos

- Todos los datos de *cuentas* de las cuentas cuyo titular es Pablo Garré.
- Nombre y apellido de los clientes que tienen préstamos en dos o más sucursales distintas.
- Nombre y apellido de los clientes con dos o más cuentas.
- Nombre y apellido de los clientes que tienen saldo superior a 50000\$ en sus cuentas.
- Nombre y apellido de los clientes que tienen saldo superior a 50000\$ en todas sus cuentas.

EBD2012_6 - Mg. Mercedes Vitturini

Otras operaciones de DML (LMD)

- Las operaciones hasta ahora vistas permiten *consultar* los datos almacenados en una Base de Datos relacional.
- El grupo de operaciones que veremos a continuación permiten *“actualizar”* el contenido de la Base de Datos, fundamentalmente por medio de las operaciones de:
 - **inserción, borrado y actualización.**

EBD2012_6 - Mg. Mercedes Vitturini

Operaciones del LMD

- **Eliminación:** consiste en borrar o eliminar tuplas de una relación.
 - Notación: $r \leftarrow r - E$.
- Ejemplo:
 - Eliminar las tuplas de corre con nro-carrera CC-02.
 $corre \leftarrow corre - \sigma_{nro-carrera = CC-02}(corre)$

EBD2012_6 - Mg. Mercedes Vitturini

Operaciones del LMD

- **Inserción:** consiste en agregar o insertar nuevas tuplas a una relación.
 - Notación: $r \leftarrow r \cup E$.
- Ejemplo:
 - Agregar el caballo “Yago” con número CB-06 de 6 años y 240 kilos.
 $caballos \leftarrow caballos \cup \{(CB-06, \text{“Yago”}, 6, 240)\}$

EBD2012_6 - Mg. Mercedes Vitturini

Operaciones de LMD

- **Modificación:** consiste en modificar datos de una o más tuplas a una relación.
 - Notación: $\delta_{A \leftarrow E}(r)$.
- Ejemplo:
 - Aumentar en 1 las edades de todos los caballos.
 $\delta_{edad \leftarrow edad + 1}(\text{caballos})$

EBD2012_6 - Mg. Mercedes Vitturini

Temas de la Clase de Hoy

- Cálculo relacional de tuplas
 - Definición formal.
 - Seguridad de las expresiones. Ejemplos
- Otros operadores de LMD
 - Operadores de actualización: eliminación, inserción y modificación
- **Bibliografía**
 - “Database System Concepts” – A. Silberschatz. Capítulo 2 (edición 2005).
 - “Database and Knowledge Base System” – J. Ullman. Capítulo 3.

EBD2012_6 - Mg. Mercedes Vitturini

Variantes de Join - Outer Join

- El **outer join** es una extensión de la operación **join** que evita la pérdida de información cuando no hay coincidencia.
- Calcula el join y luego se agregan las tuplas que no hacen match con las tuplas de la otra relación del join.
- Usa valores *nu*los o indeterminado en los atributos que no se puede determinar el valor.
- *null* significa que el valor es desconocido o no existe.

EBD2012_6 - Mg. Mercedes Vitturini

Outer Join – Ejemplo

- Relación *préstamos*

número-prest	sucursal-prest	importe
L-170	Downtown	3000
L-230	Redwood	4000
L-260	Perryridge	1700

- Relación *presta*

cliente	número-prest
Jones	L-170
Smith	L-230
Hayes	L-155

EBD2012_6 - Mg. Mercedes Vitturini

Outer Join – Ejemplo

- **Join (inner-join):** *prestamos* |><| *presta*

número-prest	sucursal	importe	cliente
L-170	Downtown	3000	Jones
L-230	Redwood	4000	Smith

- **Left-outer-Join :** *prestamos* |><| *presta*

número-prest	sucursal	importe	cliente
L-170	Downtown	3000	Jones
L-230	Redwood	4000	Smith
L-260	Perryridge	1700	<i>null</i>

EBD2012_6 - Mg. Mercedes Vitturini

Outer Join – Example

- **Right-outer-Join:** *prestamos* |><| *presta*

número-prest	sucursal	importe	cliente
L-170	Downtown	3000	Jones
L-230	Redwood	4000	Smith
L-155	<i>null</i>	<i>null</i>	Hayes

- **Full-outer-Join:** *prestamos* |><| *presta*

Número-prest	sucursal	importe	cliente
L-170	Downtown	3000	Jones
L-230	Redwood	4000	Smith
L-260	Perryridge	1700	<i>null</i>
L-155	<i>null</i>	<i>null</i>	Hayes

EBD2012_6 - Mg. Mercedes Vitturini