Universidad Simón Bolívar
Depto. de Computación y Tecnología de la Información
Sistemas Bases de Datos I (CI-3311)
Abril-Julio 2009

Restricciones de Integridad Explícitas y de Cálculo Relacional
Prof. María Esther Vidal

1. Lenguaje de la Lógica de Predicados

1.1. Vocabulario

El vocabulario de la Lógica de Predicados está compuesto por:

- Conjunto de símbolos de variables, usualmente denotadas por x, y, z, ...
- Conjunto de símbolos de constantes, usualmente denotadas por a, b, c, ...
- Conjunto de símbolos funcionales, usualmente denotadas por f, g, h, ...
- Conjunto de símbolos relacionales, usualmente denotadas por P, R, Q, ...
- Conectores lógicos.
- Cuantificadores.
- Símboloss de puntuación, es decir, paréntesis, corchetes, etc.

1.2. Sintaxis

Definición de Términos:

- Si x es un símbolo de variable, entonces x es un término.
- Si x es un símbolo de constante, entonces x es un término.
- Si f es un símbolo de función de aridad n y t_1, t_2,...,t_n son términos, entonces, $f(t_1, ..., t_n)$ es un término.

Definición de Fórmulas Bien Formadas:

- Si P es un símbolo de relación de aridad n y t_1, t_2,...,t_n son términos, entonces, $P(t_1, ..., t_n)$ es una fórmula bien formada.
- Si σ es una fórmula bien formada, $\neg\sigma$ es una fórmula bien formada.

- Si σ_i y σ_j son fórmulas bien formadas y OP es una conector lógico binario, $(\sigma_i OP \sigma_j)$ es una fórmula bien formada.
- Si σ_i y σ_j son fórmulas bien formadas y x es un símbolo variable, entonces $(\forall x \mid \sigma_i : \sigma_j)$ y $(\exists x \mid \sigma_i : \sigma_j)$ son fórmulas bien formadas.

2. Lenguage para la especificación de Restricciones Explíctas en Modelos de Datos(ER, ERE, Relacional)

2.1. Vocabulario

El vocabulario de la Lógica de Predicados está compuesto por:

- Conjunto de símbolos de variables, usualmente denotadas por x, y, z, ...
- Conjunto de símbolos de constantes, usualmente denotadas por a, b, c, ...
- Conjunto de símbolos funcionales, usualmente denotadas por f, g, h, ...
- Conjunto de símbolos de roles, usualmente denotadas por r_0, r_1, r_2, ...
- Conjunto de símbolos de atributos, usualmente denotadas por a_0, a_1, a_2, ...
- Conjunto de símbolos de predicados que denotan tipos de entidades, interrelaciones o relaciones, usualmente denotadas por P, R, Q, ...
- Conectores lógicos.
- Cuantificadores.
- Símbolos de puntuación, es decir, paréntesis, corchetes, etc.

2.2. Sintaxis

Definición de Términos:

- Si x es un símbolo de variable, entonces x es un término.
- Si x es un símbolo de constante, entonces x es un término.
- Si R es un término y a es un atributo, entonces $R.a$ es un término.
- Si R es una variable, r_i es un rol y a es un atributo, entonces $R[r_i].a$ y $R[r_i]$ son términos.
- Si f es un símbolo de función de aridad n y t_1, t_2,...,t_n son términos, entonces, $f(t_1, ..., t_n)$ es un término.

Definición de Fórmulas Bien Formadas:

- Si P es un símbolo de predicado que denota a un conjunto entidad, interrelación o relación de aridad n y t_1, t_2,...,t_n son términos, entonces, $P(t_1, ..., t_n)$ es una fórmula bien formada.
- Si σ es una fórmula bien formada, $\neg\sigma$ es una fórmula bien formada.
- Si σ_i y σ_j son fórmulas bien formadas y OP es una conector lógico binario, $(\sigma_i OP \sigma_j)$ es una fórmula bien formada.
- Si σ_i y σ_j son fórmulas bien formadas y x es un símbolo variable, entonces $(\forall x \mid \sigma_i : \sigma_j)$ y $(\exists x \mid \sigma_i : \sigma_j)$ son fórmulas bien formadas.
- Si σ_i y σ_j son fórmulas bien formadas, x es un símbolo variable, m y n son números naturales, entonces $(\exists^m x \mid \sigma_i : \sigma_j)$, $(\exists^m_n x \mid \sigma_i : \sigma_j)$, $(\exists_n x \mid \sigma_i : \sigma_j)$, $(\exists! nx \mid \sigma_i : \sigma_j)$, son fórmulas bien formadas.

2.3. Axiomas

- Axioma de Intercambio:
$$(\forall x \mid R : P) \equiv (\forall x \mid : R \Rightarrow P)$$
- Axioma de De Morgan Generalizado:
$$(\exists x \mid R : P) \equiv \neg(\forall x \mid R : \neg P)$$
- Teoremas de Intercambio:
$$(\exists x \mid R : P) \equiv (\exists x \mid : R \wedge P)$$
$$(\exists x \mid Q \wedge R : P) \equiv (\exists x \mid Q : R \wedge P)$$
- Al menos n elementos en el rango de R satisfacen Q.
$$(\exists_n x \mid R : P) \equiv n \leq (\sum x \mid R \wedge P : 1)$$
- A lo más n elementos en el rango de R satisfacen Q.
$$(\exists^n x \mid R : P) \equiv n \geq (\sum x \mid R \wedge P : 1)$$
- Entre n y m elementos en el rango de R satisfacen Q.
$$(\exists^n_m x \mid R : P) \equiv (\exists^n x \mid R : P) \wedge (\exists_m x \mid R : P)$$
- Exactamente n elementos en el rango de R satisfacen Q.
$$(\exists!_n x \mid R : P) \equiv (\exists^n_n x \mid R : P)$$

3. Cálculo relacional por tuplas

Una expresión del cálculo relacional por tuplas se define como sigue:

$$\{t_1.A_1, t_2.A_2, ..., t_n.A_n \mid COND(t_1, t_2, ..., t_n)\}$$

Donde $COND(t_1, t_2, ..., t_n)$ es una expresión del calćulo relacional que se define comos sigue. Las variables t_i deben estar asociadas a un predicado no negativo en la expresión $COND(t_1, t_2, ..., t_n)$.

3.1. Vocabulario

El vocabulario del cáculo relacional está compuesto por:

- Conjunto de símbolos de variables, usualmente denotadas por x, y, z, ...
- Conjunto de símbolos de constantes, usualmente denotadas por a, b, c, ...
- Conjunto de símbolos de funciones, usualmente denotadas por f, g, h, ...
- Conjunto de símbolos de roles, usualmente denotadas por r_0, r_1, r_2, ...
- Conjunto de símbolos de atributos, usualmente denotadas por a_0, a_1, a_2, ...
- Conjunto de símbolos de predicados que denotan tipos de entidades, interrelaciones o relaciones, usualmente denotadas por P, R, Q, ...
- Conectores lógicos.
- Cuantificadores.
- Símboloss de puntuación, es decir, paréntesis, corchetes, etc.

3.2. Sintaxis

Definición de Términos:

- Si x es un símbolo de variable, entonces x es un término.
- Si x es un símbolo de constante, entonces x es un término.
- Si R es una variable y a es un atributo, entonces $R.a$ es un término.
- Si R es una variable, r_i es un rol y a es un atributo, entonces $R[r_i].a$ y $R[r_i]$ son términos.
- Si f es un símbolos de función de aridad n y t_1, t_2,...,t_n son términos, entonces, $f(t_1, ..., t_n)$ es un término.

Definición de Fórmulas Bien Formadas:

- Si P es un símbolo de relación de aridad n y t_1, t_2,...,t_n son términos, entonces, $P(t_1, ..., t_n)$ es una fórmula bien formada.
- Si σ es una fórmula bien formada, $\neg\sigma$ es una fórmula bien formada.
- Si σ_i y σ_j son fórmulas bien formadas y OP es una conector lógico binario, $(\sigma_i OP \sigma_j)$ es una fórmula bien formada.
- Si σ_i es una fórmula bien formada y x es un símbolo variable, entonces $(\exists^m x)(\sigma_i)$ y $(\forall x)(\sigma_i)$ son fórmulas bien formadas.

3.3. Fórmulas Seguras

Una expresión E del cálculo relacional es segura si el resultado de la evaluación de la consulta correspondiente es un subconjunto finito del dominio de la misma. Es una expresión segura toda variable debe estar instanciada en un predicado no negativo.

4. Ejemplos

1. Considere el esquema presentado en la Figura **??**, donde se modela la relación que existe entre un bebé, su madre, y los dos médicos que atienden su nacimiento.

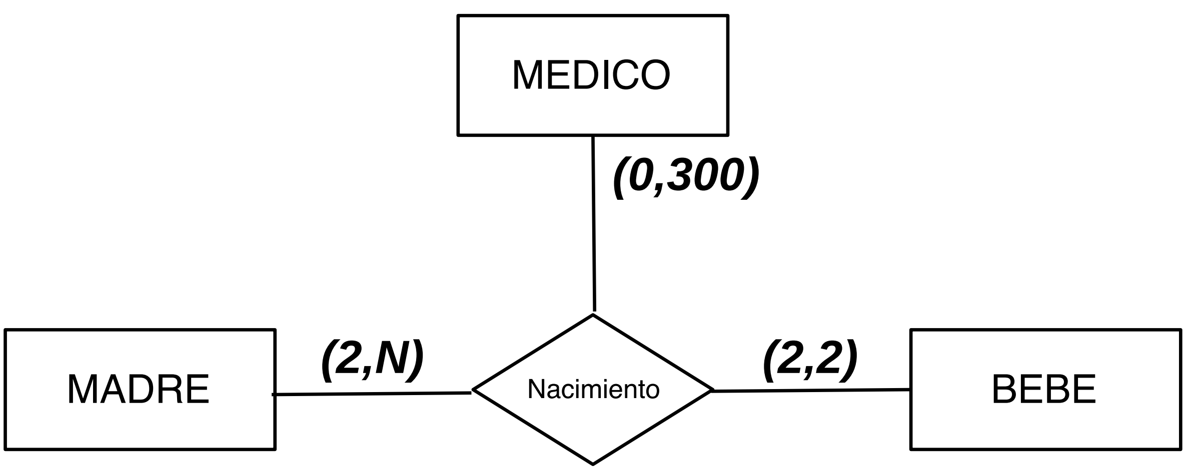

Figura 1: Nacimiento

Adicionalmente a estas restricciones implícitamente representadas en dicho diagrama, considere las siguientes restricciones explícitas. Note las expresiones de lógica utilizadas para expresar dichas restricciones:

a) Una madre puede ser atendida por el mismo médico en a lo más tres nacimientos.

$(\forall m, md \mid madre(m) \wedge medico(md) : (\exists^3 n \mid nacimiento(n) : n[madre] = m \wedge n[medico] = md))$

b) Un médico puede atender los nacimientos de a lo más 100 mujeres diferentes.

$$(\forall md \mid medico(md) : (\exists^{100} m \mid madre(m) : (\exists_1^3 n \mid nacimiento(n) : n[madre] = m \wedge n[medico] = md)))$$

c) Cada bebé tiene una única madre.

$$(\forall b \mid bebe(b) : (\exists!_1 m \mid: (\forall n \mid nacimiento(n) \wedge n[bebe] = b : n[madre] = m)))$$

Por otro lado, si tomamos en cuenta toda la información representada en el esquema y las restricciones explícitas anteriores, podemos afirmar que las siguientes afirmaciones son ciertas.

a) Todo bebé participa en la interrelación *nacimiento* en exactamente dos tripletas:

$$(\forall b \mid bebe(b) : (\exists!_2 n \mid nacimiento(n) : n[bebe] = b))$$

b) Toda madre participa como mínimo en dos tripletas de la interrelación *nacimiento*:

$$(\forall m \mid madre(m) : (\exists_2 n \mid nacimiento(n) : n[madre] = m))$$

c) Un médico va a aparecer en a lo más 300 tripletas:

$$(\forall md \mid medico(md) : (\exists^{300} n \mid nacimiento(n) : n[medico] = md))$$

Por el contrario, las siguientes afirmaciones son falsas si se considera la información representada en el esquema:

a) Existen bebés que participen exactamente una vez en la interrelación *nacimiento*:

$$(\exists b \mid bebe(b) : (\exists!_1 n \mid nacimiento(n) : n[bebe] = b))$$

b) Existen madres que no tienen hijos:

$$(\exists m \mid madre(m) : \neg(\exists n \mid nacimiento(n) : n[madre] = m))$$

c) Todo médico participa en la interrelación *nacimiento*:

$$(\forall md \mid medico(md) : (\exists n \mid nacimiento(n) : n[medico] = md))$$

2. Considere el esquema relacional de los bebedores:

FUENTES_SODA($\underline{\text{CodFS}}$,NombreFS)	BEBIDAS($\underline{\text{CodBeb}}$,NombreBeb)
BEBEDOR($\underline{\text{CI}}$,Nombre)	GUSTA($\underline{\text{CI,CodBeb}}$)
FRECUENTA($\underline{\text{CI,CodFS}}$)	VENDE($\underline{\text{CodFS,CodBeb}}$)

Las siguientes expresiones corresponden a ejemplos de cómo el cálculo relacional puede ser usado para representar consultas en bases de datos relacionales. Note que todas las expresiones son seguras, es decir, las variables *proyectadas* en la consulta, están ligadas a predicados no negados:

a) Dar las cédulas de los bebedores que les gustan bebidas que le gustan al bebedor con cédula 332.

$$\{b.CI \mid bebedor(b) \wedge (\exists g_1)(\exists g_2)(gusta(g_1) \wedge gusta(g_2) \wedge g_1.CI = b.Ci \wedge g_2.CodBeb = g_1.CodBeb \wedge g_2.CI =' 332')\}$$

b) Dar las cédulas de los bebedores que les gustan las bebidas que le gustan al bebedor con cédula 332.

$$\{b.CI \mid bebedor(b) \wedge (\forall g_1)(\neg(gusta(g_1) \wedge g_1.CI =' 332') \vee (\exists g_2)(gusta(g_2) \wedge g_2.CI = b.Ci \wedge g_2.CodBeb = g_1.CodBeb))\}$$

c) Dar las cédulas de los bebedores que les gustan sólo bebidas que le gustan al bebedor con cédula 332.

$$[\{b.CI \mid bebedor(b) \wedge (\forall g_1)(\neg(gusta(g_1) \wedge g_1.CI = b.CI) \vee (\exists g_2)(gusta(g_2) \wedge g_2.CI =' 332' \wedge g_2.CodBeb = g_1.CodBeb))\}$$

d) Dar las cédulas de los bebedores que les gustan sólo las bebidas que le gustan al bebedor con cédula 332.

$$\{b.CI \mid bebedor(b) \wedge (\forall g_1)(\neg(gusta(g_1) \wedge g_1.CI = b.CI) \vee (\exists g_2)(gusta(g_2) \wedge g_2.CI =' 332' \wedge g_2.CodBeb = g_1.CodBeb)) \wedge (\forall g_1)(\neg(gusta(g_1) \wedge g_1.CI =' 332') \vee (\exists g_2)(gusta(g_2) \wedge g_2.CI = b.Ci \wedge g_2.CodBeb = g_1.CodBeb))\}$$

e) Dar las cédulas de los bebedores que no les gustan las bebidas que le gustan al bebedor con cédula 332.

$$\{b.CI \mid bebedor(b) \wedge (\forall g_1)(\neg(gusta(g_1) \wedge g_1.CI =' 332') \vee \neg(\exists g_2)(gusta(g_2) \wedge g_2.CI = b.Ci \wedge g_2.CodBeb = g_1.CodBeb))\}$$