

## Formalización: ¿Qué dejamos fuera?

---

Sólo consideramos atributos de tipo **CDATA** y requisito **#REQUIRED**.

- Las listas de valores y atributos con requisito **#IMPLIED** o **#FIXED** pueden ser representados usando las otras alternativas.

Ejemplo:

```
<!ELEMENT book (author*)>  
<!ATTLIST book title CDATA #IMPLIED>
```

es equivalente a:

```
<!ELEMENT book (title?,author*)>  
<!ELEMENT title EMPTY>  
<!ATTLIST title valor CDATA #REQUIRED>
```

## Formalización: ¿Qué dejamos fuera?

---

- ID, IDREF y IDREFS no son un buen mecanismo para definir identificadores y hacer referencias.
  - ¿Por qué?
  - Los vamos a reemplazar por un mecanismo mejor ...

## Restricciones de integridad en XML

---

Dado: Conjunto de etiquetas  $E$  y un conjunto de atributos  $A$ .

**Clave:** Un conjunto de atributos **identifica** a los elementos con una etiqueta particular.

Dados  $\tau \in E$  y  $X \subseteq A$ , la siguiente es una clave:

$$\tau[X] \rightarrow \tau$$

## Restricciones de integridad en XML

---

Semántica: Un  $(E, A)$ -árbol  $T = (D, \lambda, \{\rho_a\}_{a \in A})$  satisface  $\tau[X] \rightarrow \tau$  si

para cada  $s_1, s_2 \in D$  tales que  $\lambda(s_1) = \lambda(s_2) = \tau$ ,

si  $\rho_a(s_1) = \rho_a(s_2)$  para cada  $a \in A$ , entonces  $s_1 = s_2$ .

Notación:  $T \models \tau[X] \rightarrow \tau$ .

Ejemplos: `book[title] → book` y `persona[rut] → persona`.

## Restricciones de integridad en XML

---

**Clave foránea:** Un conjunto de atributos hace **referencia** a la clave de los elementos con una cierta etiqueta.

Dados  $\tau_1, \tau_2 \in E$  y listas  $X, Y \subseteq A$  tales que  $|X| = |Y|$ , la siguiente es una clave foránea:

$$\tau_1[X] \subseteq_{FK} \tau_2[Y]$$

## Restricciones de integridad en XML

---

Semántica: Un  $(E, A)$ -árbol  $T = (D, \lambda, \{\rho_a\}_{a \in A})$  satisface  $\tau_1[X] \subseteq_{FK} \tau_2[Y]$  si  $T \models \tau_2[Y] \rightarrow \tau_1$  y

para cada  $s_1 \in D$  tal que  $\lambda(s_1) = \tau_1$ ,  
existe  $s_2 \in D$  tal que  $\lambda(s_2) = \tau_2$  y  $\rho_a(s_1) = \rho_a(s_2)$  para cada  $a \in A$ .

Notación:  $T \models \tau_1[X] \subseteq_{FK} \tau_2[Y]$ .

Ejemplo: `review[title]  $\subseteq_{FK}$  book[title]`.

## Claves, claves foráneas y DTDs

---

Las restricciones de integridad están asociadas a un DTD.

- El DTD indica si una restricción puede ser evaluada sobre un documento XML.

Una restricción de integridad  $\varphi$  está definida sobre un DTD

$d = (e, p, q)$  si:

- $\varphi = \tau[X] \rightarrow \tau$  y  $X \subseteq q(\tau)$ .
- $\varphi = \tau_1[X] \subseteq_{FK} \tau_2[Y]$ ,  $X \subseteq q(\tau_1)$  e  $Y \subseteq q(\tau_2)$ .

## Esquema XML completo

---

**Esquema XML:** DTD  $d$  y un conjunto  $\Sigma$  de restricciones.

Ejemplo:  $d =$

```
<!DOCTYPE db [  
  <!ELEMENT db (book*)>  
  <!ELEMENT book (author*)>  
  <!ATTLIST book  
    title CDATA #REQUIRED>  
  <!ELEMENT author EMPTY>  
  <!ATTLIST author  
    name CDATA #REQUIRED>  
>
```

$\Sigma = \{\text{book}[\text{title}] \rightarrow \text{book}, \text{author}[\text{name}] \rightarrow \text{author}\}.$

## DTDs y restricciones de integridad: Consistencia

---

Decimos que un esquema XML  $(d, \Sigma)$  es **consistente** si existe un árbol  $T$  tal que  $T \models d$  y  $T \models \varphi$  para cada  $\varphi \in \Sigma$ .

¿Cómo podemos determinar si  $(d, \Sigma)$  es consistente?

- Este problema es conocido como **análisis estático de consistencia**.
- Este es un problema no trivial. Vamos a ver su solución más adelante ...

## Restricciones de integridad relativas

---

```
<!ELEMENT db (curso*)>
<!ELEMENT curso (alumno+)>
<!ATTLIST curso
  sigla CDATA #REQUIRED>
<!ELEMENT alumno EMPTY>
<!ATTLIST alumno
  numero CDATA #REQUIRED
  nota CDATA #REQUIRED>
```

¿Cuáles de las siguientes restricciones son correctas?

```
curso[sigla] → curso
alumno[numero] → alumno
```

## Claves relativas

---

Dado: Conjunto de etiquetas  $E$  y un conjunto de atributos  $A$ .

**Clave relativa:** Un conjunto de atributos identifica a los elementos con una etiqueta particular, **siempre y cuando sean descendientes de un elemento con una cierta etiqueta.**

Dados  $\tau, \tau_1 \in E$  y  $X \subseteq A$ , la siguiente es una clave relativa:

$$\tau(\tau_1[X]) \rightarrow \tau_1$$

## Claves relativas: Semántica

---

Un  $(E, A)$ -árbol  $T = (D, \lambda, \{\rho_a\}_{a \in A})$  satisface  $\tau(\tau_1[X] \rightarrow \tau_1)$  si

para cada  $s \in D$  tal que  $\lambda(s) = \tau$ ,

para cada  $s_1, s_2 \in D$  tales que  $\lambda(s_1) = \lambda(s_2) = \tau_1$  y

$s_1, s_2$  son descendientes de  $s$  en  $T$ ,

si  $\rho_a(s_1) = \rho_a(s_2)$  para cada  $a \in A$ , entonces  $s_1 = s_2$ .

Notación:  $T \models \tau(\tau_1[X] \rightarrow \tau_1)$ .

Ejemplo: `curso(alumno[numero] → alumno)`.

## Claves foráneas relativas

---

**Clave foránea relativa:** Un conjunto de atributos hace referencia a la clave de los elementos con una cierta etiqueta, **siempre y cuando estos elementos sean descendientes de un elemento de un cierto tipo.**

Dados  $\tau, \tau_1, \tau_2 \in E$  y listas  $X, Y \subseteq A$  tales que  $|X| = |Y|$ , la siguiente es una clave foránea relativa:

$$\tau(\tau_1[X] \subseteq_{FK} \tau_2[Y])$$

## Restricciones de integridad en XML

---

Semántica: Un  $(E, A)$ -árbol  $T = (D, \lambda, \{\rho_a\}_{a \in A})$  satisface  $\tau(\tau_1[X] \subseteq_{FK} \tau_2[Y])$  si  $T \models \tau(\tau_2[Y] \rightarrow \tau_2)$  y

para cada  $s \in D$  tal que  $\lambda(s) = \tau$ ,

para cada  $s_1 \in D$  tal que  $\lambda(s_1) = \tau_1$  y

$s_1$  es un descendiente de  $s$  en  $T$

existe  $s_2 \in D$  tal que  $\lambda(s_2) = \tau_2$ ,  $s_2$  es un descendiente de  $s$  en  $T$  y  $\rho_a(s_1) = \rho_a(s_2)$  para cada  $a \in A$ .

Notación:  $T \models \tau(\tau_1[X] \subseteq_{FK} \tau_2[Y])$ .

Ejercicio: De un ejemplo natural de una clave foránea relativa.

## Restricciones relativas y DTDs

---

Al igual que en el caso las restricciones de integridad *absolutas*, las restricciones de integridad relativas están asociadas a un DTD.

Notación: Una etiqueta  $\tau_1$  es descendiente de una etiqueta  $\tau_2$  en un DTD  $d$  si existe un árbol  $T$  y elementos  $s_1, s_2$  en  $T$  tales que  $T \models d$ ,  $\lambda(s_1) = \tau_1$ ,  $\lambda(s_2) = \tau_2$  y  $s_1$  es descendiente de  $s_2$  en  $T$ .

## Restricciones relativas y DTDs

---

Una restricción de integridad relativa  $\varphi$  está definida sobre un DTD  $d = (e, p, q)$  si:

- $\varphi = \tau(\tau_1[X] \rightarrow \tau_1)$ ,  $X \subseteq q(\tau_1)$  y  $\tau_1$  es descendiente de  $\tau$  en  $d$ .
- $\varphi = \tau(\tau_1[X] \subseteq_{FK} \tau_2[Y])$ ,  $X \subseteq q(\tau_1)$ ,  $Y \subseteq q(\tau_2)$  y tanto  $\tau_1$  como  $\tau_2$  son descendientes de  $\tau$  en  $d$ .