

ACTIVIDADES VALIDACIÓN DE DOCUMENTOS XML A TRAVÉS DE DTD

En el propio documento:

Sintaxis:

```
<!DOCTYPE raíz [...códigoDTD...]>
```

Ejemplo:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE persona [
    <!ELEMENT persona (nombre)>
    <!ELEMENT nombre (#PCDATA)>
]>
<persona>
    <nombre>Antonio</nombre>
</persona>
```

En un documento externo privado:

Sintaxis:

```
<!DOCTYPE raíz SYSTEM "ruta_URL_al_DTD">
```

Ejemplo:

Documento ciudades.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ciudades SYSTEM "ciudades.dtd">

<ciudades>
    <ciudad pais="Italia">Roma</ciudad>
    <ciudad pais="Francia">París</ciudad>
    <ciudad pais="Alemania">Berlín</ciudad>
    <ciudad pais="">Viena</ciudad>
</ciudades>
```

Documento ciudades.dtd

```
<!ELEMENT ciudades (ciudad)*>
<!ELEMENT ciudad (#PCDATA)>
<!ATTLIST ciudad pais CDATA #REQUIRED>
```

En un documento externo público:

```
<!DOCTYPE raíz PUBLIC "nombreDTD" "DTD_URL">
```

Definiciones en un DTD

En un código DTD (tanto externo como interno) se pueden definir:

- Los elementos que se pueden utilizar en un documento XML. En esta definición se indica además que pueden contener dichos elementos.
- Los atributos que pueden poseer los elementos. Además incluso indicando sus posibles valores válidos.
- Las entidades que puede utilizar el documento XML.

Definición de elementos mediante DTD:

<!ELEMENT nombre tipo>

Tipos de contenido en los elementos:

EMPTY

El elemento no puede tener contenido

Contenido concreto

Dicho elemento se indica entre paréntesis

Nombres de elementos **<!ELEMENT persona (nombre)>**

Contenido de texto **<!ELEMENT nombre (#PCDATA)>**

ANY

Permite cualquier contenido en el elemento, sin restricciones de ningún tipo. Es decir puede contener texto, otro tipo de datos y cualquier etiqueta. Además puede tener atributos.

Ejemplo:

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<!DOCTYPE persona [
```

```
  <!ELEMENT persona ANY>
```

```
  <!ELEMENT nombre (#PCDATA)>
```

```
  <!ELEMENT ciudad (#PCDATA)>
```

```
<persona>
```

```
  <nombre>Mario</nombre> vive en <ciudad>Madrid</ciudad>.
```

```
</persona>
```

Secuencias

Ejemplo:

```
<!ELEMENT persona (nombre, apellidos, edad)>
  <!ELEMENT nombre (#PCDATA)>
  <!ELEMENT apellidos (#PCDATA)>
  <!ELEMENT edad (#PCDATA)>
```

Elecciones u opciones

Ejemplo:

```
<!ELEMENT artículo (id | serie)>
  <!ELEMENT id (#PCDATA)>
  <!ELEMENT serie (#PCDATA)>
```

Combinaciones

```
<!ELEMENT coordenada ((longitud, latitud) | coordUniversal)>
<!ELEMENT longitud (#PCDATA)>
<!ELEMENT latitud (#PCDATA)>
<!ELEMENT coordUniversal (#PCDATA)>
```

Ejercicio 1:

Guarde la declaración anterior en un documento .dtd y diseñe un documento XML bien formado y válido según esa declaración.

Cardinalidad

La cardinalidad es el número de veces que puede aparecer un determinado contenido en un elemento. Se realiza mediante estos símbolos:

- ? Contenido opcional, puede aparecer (una sola vez) o no aparecer
- * Contenido opcional y repetible. Es decir puede no aparecer y puede incluso aparecer varias veces
- + Contenido obligatorio y repetible. Tiene que aparecer e incluso puede aparecer varias veces

Ejemplo 1:

```
<!ELEMENT película (titulo, direccion+, argumento?, actor*)>
```

Ejemplo 2:

```
<!ELEMENT polígono ((coordX,coordY)+ | nombre)>
<!ELEMENT coordX (#PCDATA)>
<!ELEMENT coordY (#PCDATA)>
<!ELEMENT nombre (#PCDATA)>
```

Ejercicio 2:

Crear un documento XML bien formado y válido para cada una de las declaraciones anteriores (Ejemplo 1 y Ejemplo 2)

Declaración de atributos mediante DTD

<!ATTLIST elemento nombreAtributo tipo presencia valorPorDefecto>

Donde:

- elemento. Es el nombre del elemento que podrá utilizar el atributo
- nombreAtributo. Es el identificador del atributo que estamos declarando (y que debe de cumplir las reglas de identificadores de XML)
- tipo. Es el tipo de valores que podemos asignar al atributo
- presencia. Indica las características de los valores que puede tomar el atributo: si es obligatorio, si hay valor por defecto,...
- valorPorDefecto. Permite especificar un valor que el atributo tomará si no especifica otro explícitamente en el documentoXML.

Ejemplos:

```
<!ATTLIST persona nacionalidad CDATA>
```

(CDATA sirve para indicar que el atributo contiene texto. A diferencia de los datos PCDATA de los elementos, los CDATA admiten cualquier carácter del tipo que sea)

```
<!ATTLIST persona nacionalidad CDATA "Española">
```

valores fijos

Se utiliza el término #FIXED

```
<!ATTLIST persona nacionalidad CDATA #FIXED "Española">
```

El atributo nacionalidad no podrá tomar ningún valor que no sea el valor Española, impidiendo tomar otra nacionalidad.

valores requeridos

En este caso se usa la palabra #REQUIRED indicando con ello que siempre hay que dar valor al atributo. Ejemplo:

```
<!ATTLIST persona nacionalidad CDATA #REQUIRED>
```

Un documento XML que utilice el elemento persona deberá especificar obligatoriamente la nacionalidad.

Obviamente, el uso de #REQUIRED no permite indicar un valor por defecto, al no poder dejarse sin especificar el atributo.

valor opcional

La palabra #IMPLIED especificada en el atributo indicaría que dicho atributo puede quedarse sin valor; es decir no posee valor por defecto, pero puede quedarse sin especificar (quedaría nulo por tanto).

```
<!ATTLIST persona nacionalidad CDATA #IMPLIED>
```

En el ejemplo, el atributo nacionalidad no es obligatorio especificarle, puede quedar sin valor. Nuevamente en este caso no se puede especificar un valor por defecto (sería absurdo).

enumeraciones

Se trata de atributos a los que se indica una serie de posibles valores. El valor del atributo debe coincidir con alguno de la lista. Los valores posibles se indican entre paréntesis separados por el símbolo |.

Ejemplo:

```
<!ATTLIST persona sexo (Hombre | Mujer) #REQUIRED >
```

Las personas sólo podrán especificar como sexo “Hombre” o “Mujer” y nada más:

```
<persona sexo="Varón">Javier Ruiz</persona>
```

declaración de varios atributos en la misma etiqueta

Se usa muy habitualmente para indicar de forma cómoda todos los atributos de un determinado elemento:

```
<!ELEMENT directorio (persona)+>  
<!ELEMENT persona (#PCDATA)>  
<!--ATTLIST persona  nacionalidad CDATA "Española"  
      sexo (Hombre | Mujer) #IMPLIED  
      id ID #REQUIRED-->
```

Ejercicio 3:

Para esta declaración con varios atributos cree un documento .xml bien formado y válido.

Ejercicio 4:

Dado el siguiente árbol cuya raíz es cartelera, diseñe una DTD que lo valide, teniendo en cuenta que:

- El título original de una película sólo aparecerá cuando la película no sea española.
- Es posible que en un momento dado una película esté pendiente de clasificación. En caso de que esté clasificada siempre deberá indicar los años para los que se recomienda: tp (todos los públicos), 8, 12, 16 o 18.
- No siempre existe una web con la información de la película.
- Se quiere guardar información sobre el fichero gráfico que contiene el cartel de la película. Este fichero no siempre está disponible.
- En caso de que no se proporcione el año de una película se asumirá que es el 2003.
- En el reparto deberá aparecer un actor como mínimo.

<cartelera>

<película código="p1" duración="152" año="2002">
<título>AQUELLAS JUERGAS UNIVERSITARIAS</título>
<título_original>Old School</título_original>
<nacionalidad>Estados Unidos</nacionalidad>
<género>Comedia</género>
<clasificación edad="tp"/>
<sinopsis>

Mitch, Frank y Beanie son tres amigos treintañeros cuyas vidas no son exactamente lo que esperaban. Mitch tiene una novia ninfómana que se mete en la cama con el primero que agarra.

Frank se ha casado y su matrimonio nada

tiene que ver con las juergas salvajes que organizaban años atrás. Y Beanie es

un padre de familia que se muere por recuperar su alocada juventud. Pero las cosas cambian

cuando Beanie sugiere que creen su propia fraternidad, en la nueva casa que Mitch tiene junto

al campus de la universidad. Una ocasión para revivir tiempos gloriosos, hacer nuevos amigos y

de volver a sus viejas, salvajes y desmadradas juergas de estudiantes.

</sinopsis>

<director>Todd Philips</director>

<reparto>

<actor>Luke Wilson</actor>

<actor>Will Farrel</actor>

<actor>Vince Vaughn</actor>

</reparto>

<web><http://www.uip.es></web>

<cartel>caratulas/Aquellas juergas.jpg</cartel>

</película>

<película código="p17" duración="06">
<título>EL ORO DE MOSCÚ</título>
<nacionalidad>España</nacionalidad>
<género>Comedia</género>
<sin_clasificar/>
<sinopsis>

Por una extraña coincidencia del destino, alguien recibe una información extraconfidencial de un anciano en sus últimos segundos de vida: el secreto mejor guardado de la Historia. El receptor, un trabajador de hospital, se lo comunica secretamente a un supuesto amigo. Ambos inician una aventura rocambolesca y llena de misterio. Ante la inutilidad de sus intentos y muy a su pesar, tienen que recurrir a otras personas que así mismo van cayendo en el pozo sin fondo que conlleva descifrar el enigma.

</sinopsis>

<director>Jesús Bonilla</director>

<reparto>

<actor>Jesús Bonilla</actor>

<actor>Santiago Segura</actor>

<actor>Alfredo Landa</actor>

<actor>Concha Velasco</actor>

<actor>Antonio Resines</actor>

<actor>Gabino Diego, María Barranco</actor>

<actor>María Barranco</actor>

</reparto>

</película>

</cartelera>

Referencias:

<http://xml-copy-editor.sourceforge.net/>

siguiendo este enlace instale XML COPY EDITOR