

XML XPATH

LIMA UD3 – Tema 1

IES Plurilingüe Antón Losada Diéguez

Adrián Fernández González



Tabla de contenido

1. Introducción.....	2
2. Nodo.....	2
3. Sintaxis básica	2
3.1. Seleccionar raíz	2
3.2. Seleccionar descendiente	2
3.3. Seleccionar cualquier descendiente.....	3
3.4. Seleccionar atributos	3
4. Predicados	3
4.1. Comodín.....	3
XML de referencia.....	4

XML XPath

1. Introducción

XPath (*XML Path Language*) es un lenguaje que permite procesar documentos XML en base a expresiones.

XPath se basa en la estructura jerárquica de los elementos XML para acceder a ellos como si fuese una estructura de carpetas.

Normalmente, este lenguaje se utiliza desde un lenguaje de alto nivel para un acceso rápido a ciertos elementos que de otra forma requerirían mucho más trabajo.

Este lenguaje solo puede utilizarse si se conoce la estructura del XML a manipular y suele usarse para seleccionar cosas en concreto más que recorridos secuenciales.

2. Nodo

Un nodo es todo lo que se puede representar en XML. Un elemento, texto, atributo, comentario, namespace... Todo ello es accesible de una u otra forma con XPath.

Lo único que no se considera nodo son las declaraciones DOCTYPE y la propia cabecera xml.

Un nodo especial es el nodo raíz, aquel que aúna a los demás elementos y que suele ser el elemento en el que se empezará.

3. Sintaxis básica

La sintaxis es muy similar a un acceso a carpetas, usando los caracteres '.', '/' y '@'.

Cuando se selecciona, se puede obtener tanto un único elemento o un conjunto de ellos, dependiendo de lo que concuerde con el selector.

3.1. Seleccionar raíz

Para seleccionar el nodo raíz hay que poner / y el nombre del elemento raíz.

```
/menu
```

En este ejemplo se selecciona el elemento menú, raíz del documento.

3.2. Seleccionar descendiente

Para seleccionar un descendiente de un elemento, se utiliza la / y el nombre de cada hijo hasta el elemento.

```
/menu/primeros/primero
```

En este ejemplo se obtienen todos los elementos primero que son hijos del elemento primeros que es a su vez hijo de menú.

3.3. Seleccionar cualquier descendiente

Para obtener todos los elementos descendientes de uno, independientemente de que tengan el mismo padre o no, se usa '/' y el nombre del elemento.

```
/menu//bebida
```

En este caso, se obtienen todas las bebidas del menú, tanto los que se encuentran en bebidas como la que se encuentra en el entrante.

3.4. Seleccionar atributos

Para obtener los atributos de un elemento o elementos, se accede a él y se añade '/'@' y el nombre del atributo.

```
/menu/bebidas/bebida/@alcohol
```

En este caso, se obtiene el atributo alcohol de todos los elementos bebida dentro de bebidas.

4. Predicados

Los predicados son expresiones más complejas que permiten seleccionar un elemento concreto entre sus hermanos.

Con el uso de '[' se escoge un elemento determinado mediante un número o combinado con funciones como last() para obtener el último.

```
/menu/primeros/primero[1]
```

Así en este ejemplo se obtiene el primer plato primero dentro de primeros, dentro de menú. Nótese que **los índices empiezan en 1** y no en 0 como en otros lenguajes.

```
/menu/primeros/primero[last()-1]
```

En este ejemplo se obtiene el penúltimo elemento primero.

4.1. Comodín

Para obtener todos los elementos o atributos hijos de un elemento, se puede usar el comodín '*'.

```
/menu/bebidas/bebida/@*
```

En este ejemplo, se obtienen todos los atributos de los elementos bebida dentro de bebidas.

XML de referencia

```
<menu>
  <entrantes>
    <entrante></entrante>
    <entrante></entrante>
    <bebida alcohol=""></bebida>
  </entrantes>
  <primeros>
    <primero></primero>
    <primero></primero>
    <sorbete alcohol=""></sorbete>
  </primeros>
  <segundos>
    <segundo tipo=""></segundo>
  </segundos>
  <postres>
    <helado sabor=""></helado>
    <tarta></tarta>
  </postres>
  <bebidas>
    <bebida alcohol=""></bebida>
    <bebida alcohol=""></bebida>
    <cafe></cafe>
  </bebidas>
</menu>
```