

Úvod do XPath

(zdroj: <http://w3schools.com>)

Xpath je jazyk určený pre hľadanie informácií v XML dokumentoch.

Čo je XPath?

- XPath je syntax pre definovanie častí XML dokumentu
 - XPath používa cestné výrazy pre navigáciu v XML dokumentoch
 - XPath obsahuje knižnicu štandardných funkcií
 - XPath je podstatným elementom (prvkom) v XSLT
 - XPath je W3C odporúčaním (W3C recommendation)
-

XPath cestné výrazy

XPath používa cestné výrazy na selektovanie uzlov alebo sád uzlov v XML dokumentoch. Tieto XPath cestné výrazy sú veľmi podobné výrazom používaným na definovanie ciest k adresárom alebo k súborom v klasickom súborovom systéme.

XPath štandardné funkcie

XPath zahŕňa viac ako 100 vstavaných funkcií, ktoré dokážu pracovať s hodnotami reťazcov, s numerickými hodnotami a s Boolean hodnotami. Táto sada funkcií obsahuje tiež funkcie, ktoré dokážu porovnávať dátumy a časy, vykonať uzlové, QName a sekvenčné manipulácie atď..

XPath sa používa v XSLT

XPath je podstatným prvkom v XSLT štandarde. Bez znalosti XPath nie sme schopný vytvárať XSLT dokumenty.

XPATH je W3C odporúčaním (W3C recommendation)

XPath sa stal W3C odporúčaním (W3C recommendation) 16. novembra 1999. XPath bol navrhnutý pre použitie v XSLT, XPointer a v iných XML parsovacích programoch.

XPath uzly

XPath terminológia

Uzly

V XPath poznáme sedem druhov uzlov: element, atribút, text, priestor mien, spracúvajúca inštrukcia, komentár a uzly dokumentu.

XML dokumenty sú spracovávané ako stromy uzlov. Najvyšší element stromu sa nazýva koreňovým elementom.

Použitie uvedenej terminológie ukážeme v nasledujúcom XML dokumente:

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<bookstore>
  <book>
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
</bookstore>
```

Príklady uzlov v uvedenom XML dokumente:

<bookstore> (koreňový elementový uzol)

<author>J K. Rowling</author> (elementový uzol)

lang="en" (atribútový uzol)

Atomické hodnoty

sú uzly, ktoré nemajú žiadnych rodičov ani žiadne deti.

Príklady atomických hodnôt:

J K. Rowling

"en"

Položky

sú atomické hodnoty alebo uzly.

Vzťahy uzlov

Vzťahy medzi uzlami sú jazykom XPath chápané v úplnej zhode s hierarchickými vzťahmi elemntov v XML dokumente (rodičovské, detské a súrodenecké uzly, predkovia a potomkovia uzlov).

XPath syntax

XPath používa cestné výrazy na selektovanie uzlov alebo sád uzlov v XML dokumentoch. Uzol je selektovaný pomocou sledovania cesty k nemu alebo pomocou krokov, ktorými sa dostaneme k nemu.

Ukázkový XML dokument

V nasledujúcich príkladoch budeme používať nasledujúci XML dokument:

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<bookstore>

<book>
  <title lang="eng">Harry Potter</title>
  <price>29.99</price>
</book>

<book>
  <title lang="eng">Learning XML</title>
  <price>39.95</price>
</book>

</bookstore>
```

Selektovanie uzlov

Často používané cestné výrazy sú zobrazené v nasledujúcej tabuľke (tabuľka s anglickým popisom je z <http://w3schools.com>)

Expression	Description
<i>nodename</i>	Selects all child nodes of the named node
/	Selects from the root node
//	Selects nodes in the document from the current node that match the selection no matter where they are
.	Selects the current node
..	Selects the parent of the current node
@	Selects attributes

Zoznam niekoľkých príkladov cestných výrazov vzťahujúcich sa k uvedenému XML dokumentu aj s výsledkami týchto výrazov (tabuľka s anglickým popisom je z <http://w3schools.com>):

Path Expression	Result
bookstore	Selects all the child nodes of the bookstore element
/bookstore	Selects the root element bookstore Note: If the path starts with a slash (/) it always represents an absolute path to an element!
bookstore/book	Selects all book elements that are children of bookstore
//book	Selects all book elements no matter where they are in the document
bookstore//book	Selects all book elements that are descendant of the bookstore element, no matter where they are under the bookstore element
//@lang	Selects all attributes that are named lang

Predikáty (predikát - prívlastok vypovedajúca o subjekte alebo premennej) sa používajú na nájdenie **špecifického uzla** alebo **uzla, ktorý obsahuje špecifickú hodnotu**. **Predikáty** sú vždy vložené **v hranatých zátvorkách**.

Zoznam niekoľkých príkladov cestných výrazov s predikátmi, vzťahujúcimi sa k uvedenému XML dokumentu aj s výsledkami týchto výrazov (tabuľka s anglickým popisom je z <http://w3schools.com>):

Path Expression	Result
/bookstore/book[1]	Selects the first book element that is the child of the bookstore element. Note: IE5 and later has implemented that [0] should be the first node, but according to the W3C standard it should have been [1]!!
/bookstore/book[last()]	Selects the last book element that is the child of the bookstore element
/bookstore/book[last()-1]	Selects the last but one book element that is the child of the bookstore element
/bookstore/book[position()<3]	Selects the first two book elements that are children of the bookstore element
//title[@lang]	Selects all the title elements that have an attribute named lang
//title[@lang='eng']	Selects all the title elements that have an attribute named lang with a value of 'eng'
/bookstore/book[price>35.00]	Selects all the book elements of the bookstore element that have a price element with a value greater than 35.00
/bookstore/book[price>35.00]/title	Selects all the title elements of the book elements of the bookstore element that have a price element with a value greater than 35.00

Selektovanie neznámych uzlov

XPath zástupné znaky (wildcards) môžu byť použité na selektovanie neznámych XML elementov. V nasledujúcej tabuľke je uvedený zoznam XPath zástupných znakov (tabuľka s anglickým popisom je z <http://w3schools.com>):

Wildcard	Description
*	Matches any element node
@*	Matches any attribute node
node()	Matches any node of any kind

Zoznam niekoľkých príkladov cestných výrazov so zástupnými znakmi, vzťahujúcich sa k uvedenému XML dokumentu aj s výsledkami týchto výrazov (tabuľka s anglickým popisom je z <http://w3schools.com>):

Path Expression	Result
/bookstore/*	Selects all the child nodes of the bookstore element
//*	Selects all elements in the document
//title[@*]	Selects all title elements which have any attribute

Selektovanie viacerých ciest Selecting Several Paths

Použitím operátora | v XPath výraze môžeme vyselektovať viaceré cesty.

Zoznam niekoľkých príkladov cestných výrazov používajúcich operátor | a vzťahujúcich sa k uvedenému XML dokumentu aj s výsledkami týchto výrazov (tabuľka s anglickým popisom je z <http://w3schools.com>):

Path Expression	Result
//book/title //book/price	Selects all the title AND price elements of all book elements
//title //price	Selects all the title AND price elements in the document
/bookstore/book/title //price	Selects all the title elements of the book element of the bookstore element AND all the price elements in the document

XPath osi

Ukázkový XML dokument

V nasledujúcich príkladoch budeme používať nasledujúci XML dokument:

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<bookstore>

<book>
  <title lang="eng">Harry Potter</title>
  <price>29.99</price>
</book>

<book>
  <title lang="eng">Learning XML</title>
  <price>39.95</price>
</book>

</bookstore>
```

XPath osi

Os definuje sadu uzlov vzťahujúcu sa na aktuálny uzol.

V nasledujúcej tabuľke je uvedený zoznam XPath osí (tabuľka s anglickým popisom je z <http://w3schools.com>):

AxisName	Result
ancestor	Selects all ancestors (parent, grandparent, etc.) of the current node
ancestor-or-self	Selects all ancestors (parent, grandparent, etc.) of the current node and the current node itself
attribute	Selects all attributes of the current node
child	Selects all children of the current node
descendant	Selects all descendants (children, grandchildren, etc.) of the current node
descendant-or-self	Selects all descendants (children, grandchildren, etc.) of the current node and the current node itself
following	Selects everything in the document after the closing tag of the current node
following-sibling	Selects all siblings after the current node
namespace	Selects all namespace nodes of the current node
parent	Selects the parent of the current node
preceding	Selects everything in the document that is before the start tag of the current node
preceding-sibling	Selects all siblings before the current node
self	Selects the current node

Lokalizačný cestný výraz

Lokalizačná cesta môže byť absolútna alebo relatívna.

Absolútna lokalizačná cesta začína znakom /, relatívna lokalizačná cesta týmto znakom nezačína. V oboch prípadoch však lokalizačná cesta pozostáva z jedného alebo viacerých krokov, pričom každý z nich je oddelený znakom /.

Absolútna lokalizačná cesta:

/step/step/...

Relatívna lokalizačná cesta:

step/step/...

Každý krok je vyhodnotený oproti uzlom v aktuálnej sade.

Krok pozostáva z/zo:

- osi (definuje stromový vzťah medzi vybratými uzlami a aktuálnym uzlom) an axis
- uzlového testu (identifikuje uzol vo vnútri osi)
- žiadneho alebo viacerých predikátov (pre ďalšie spresnenie vyselektovanej uzlovej sady)

Syntax lokalizačného kroku je nasledujúca:

axisname::nodetest[predicate]

Príklady použitia lokalizačných cestných výrazov (tabuľka s anglickým popisom je z <http://w3schools.com>):

Example	Result
child::book	Selects all book nodes that are children of the current node
attribute::lang	Selects the lang attribute of the current node
child::*	Selects all children of the current node
attribute::*	Selects all attributes of the current node
child::text()	Selects all text child nodes of the current node
child::node()	Selects all child nodes of the current node
descendant::book	Selects all book descendants of the current node
ancestor::book	Selects all book ancestors of the current node
ancestor-or-self::book	Selects all book ancestors of the current node - and the current as well if it is a book node
child::* / child::price	Selects all price grandchildren of the current node