XPath

Šta je XPath

jezik za označavanje delova XML dokumenta

- zasniva se na konceptu navigacije kroz stablo dokumenta
- ima biblioteku standardnih funkcija
- neophodan za korišćenje XSLT
- W3C standard

XPath izrazi

- izraz je namenjen za označavanje čvora ili skupa čvorova u dokumentu
- ovi izrazi liče na izraze za rad sa fajl-sistemom
 - fajl-sistem
 - C:\windows\fonts\arial.ttf
 - \windows\fonts\arial.ttf
 - ..\..\windows\fonts\arial.ttf

Standardne XPath funkcije

- XPath sadrži oko 100 ugrađenih funkcija
 - operacije sa stringovima
 - operacije numeričkim vrednostima
 - operacije sa datumom i vremenom
 - manipulacija čvorovima
 - manipulacija sekvencama

—

XPath čvorovi

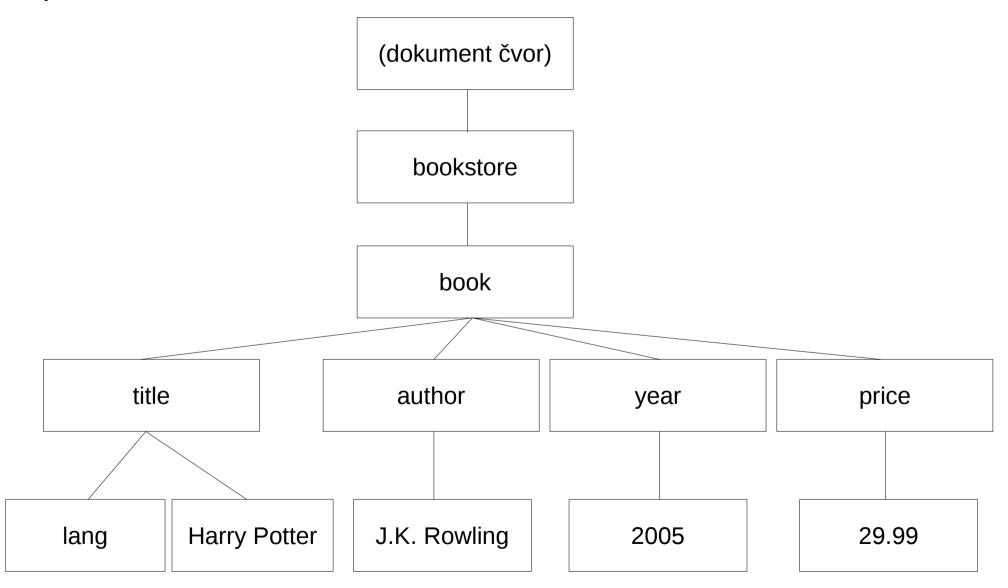
- XML dokument se, sa stanovišta XPath-a, posmatra kao stablo
- više tipova čvorova
 - element
 - atribut
 - tekst
 - namespace
 - procesna instrukcija
 - komentar
 - dokument (koren)

XPath čvorovi

primer dokument čvor <?xml version="1.0"?> <bookstore> atribut čvor <book> <title lang="en">Harry Potter</title> <author>J. K. Rowling</author> <year>2005 <price>29.99</price> element čvor </book> </bookstore>

XPath čvorovi

primer



Odnosi među čvorovima

- roditelj (book za title)
- dete (title za book)
- brat/sestra (title i author)
- predak (bookstore za title)
- naslednik (title za bookstore)

book

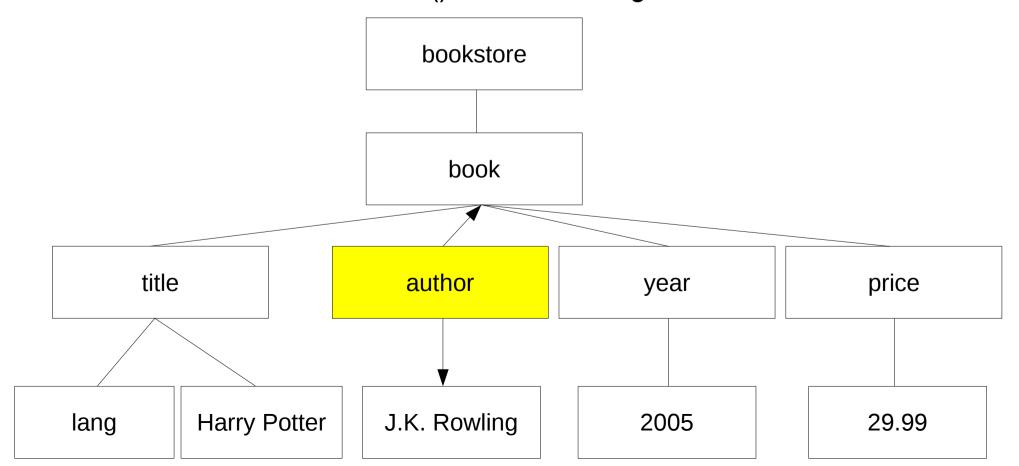
title author year price

lang Harry Potter J.K. Rowling 2005 29.99

hookstore

Navigacija kroz stablo dokumenta

- koncept tekućeg čvora = "gde se trenutno nalazim"
- ako je tekući čvor author
 - u smeru roditelj nalazi se book
 - u smeru dete nalazi se text()="J.K. Rowling"



Navigacija kroz stablo dokumenta

primer dokumenta

```
<?xml version="1.0"?>
<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 putanja

- sastoji se iz više koraka razdvojenih znakom /
- apsolutna putanja: počinje znakom /
 - /korak/korak/...
 - kreće uvek od dokument-čvora
- relativna putanja: ne počinje znakom /
 - korak/korak/...

 svaki korak se izračunava u odnosu na čvorove u tekućem skupu čvorova (node-set)

XPath putanja

- korak putanje se sastoji iz
 - ose kretanja
 - testa čvora
 - nula ili više predikata

opšti oblik izraza za korak je

```
osa::test[predikat]
```

osa definiše skup čvorova u odnosu na tekući čvor

```
ancestor svi preci
```

ancestor-or-self svi preci ili sam čvor

attribute svi atributi

child sva deca

descendantsvi potomci

descendant-or-selfsvi potomci ili sam čvor

following sve u dokumentu iza krajnjeg taga tekućeg

following-sibling sva braća posle tekućeg

namespace svi namespace čvorovi tekućeg

parent roditelj

preceding sve u dokumentu iza početnog taga tekućeg

preceding-sibling sva braća pre tekućeg

self tekući čvor

• primeri

- child::book
 - svi book čvorovi koji su deca tekućeg čvora
- attribute::lang
 - atribut lang tekućeg čvora (može biti najviše jedan)
- attribute::*
 - svi atributi tekućeg čvora
- child::text()
 - svi tekstualni čvorovi koji su deca tekućeg čvora

primeri

- descendant::book
 - svi book naslednici tekućeg čvora
- ancestor::book
 - svi book preci tekućeg čvora
- ancestor-or-self::book
 - svi book preci tekućeg čvora i tekući čvor ako je on book
- child::*/child::price
 - svi price unuci tekućeg čvora

XPath putanja

skraćeno pisanje

- ako se osa ne navede podrazumeva se child::self:: se može pisati kao tačka .
- parent:: se može pisati kao dve tačke . .
- descendant:: se može pisati kao dve kose crte //
- attribute:: se može pisati kao @
- following-sibling:: se može pisati kao . . /

• ako se osa ne navede, podrazumeva se child::

primeri

- ako se umesto ose navede @, podrazumeva se attribute::
- primeri

```
- child::title/attribute::lang ↔ title/@lang
```

- ako se umesto ose navede //, podrazumeva se descendant::
- primeri
 - descendant::title ← //title

- koriste se za pronalaženje čvora koji zadovoljava dati uslov
 - osnovna primena: filtriranje čvorova
 - primer: od svih title čvorova, izdvoj one čiji je tekstualni sadržaj duži od 20 znakova
- uvek se pišu unutar uglastih zagrada [...]

primeri

- //title[@lang]
 - svi elementi title koji imaju atribut lang
- //title[@lang='eng']
 - svi title elementi koji imaju atribut lang sa vrednošću 'eng'
- /bookstore/book[price > 35.00]
 - svi book elementi deca bookstore korena koji imaju podelement price sa vrednošću većom od 35.00
- /bookstore/book[price > 35.00]/title
 - svi title elementi koji pripadaju book elementima koji pripadaju bookstore korenu koji imaju price element sa vrednošću većom of 35.00
- /bookstore/book[price > 35.00]/price
 - svi price elementi koji pripadaju book elementima koji pripadaju bookstore korenu koji imaju price element sa vrednošću većom of 35.00

• [position() = X] se može pisati kao [X]

- primeri
 - /bookstore/book[1]
 - prvi book element koji je dete bookstore korena
 - /bookstore/book[last()]
 - poslednji book element koji je dete bookstore korena
 - /bookstore/book[last()-1]
 - pretposlednji book element koji je dete bookstore korena
 - /bookstore/book[position()<3]</pre>
 - prva dva book elementa koji su deca bookstore korena

XPath operatori

Operator	Opis	Primer	Rezultat
1	unija dva skupa čvorova	//book //cd	skup čvorova sa svim book i cd elementima
+	sabiranje	6 + 4	10
-	oduzimanje	6-4	2
*	množenje	6*4	24
div	deljenje	6 div 4	1
mod	ostatak pri deljenju	6 mod 4	2
=	jednako	price=9.80	true ako je price=9.80, inače false
!=	nejednako	price!=9.80	
<	manje	price<9.80	
<=	manje ili jednako	price<=9.80	
>	veće	price>9.80	
>=	veće ili jednako	price>=9.80	
or	disjunkcija	price=9.80 or price=9.70	
and	konjunkcija	price>9.00 and price<9.90	

Biblioteka standardnih funkcija

core functions: 27 funkcija

```
- boolean()
                               - normalize-space()
- ceiling()
                               - not()
- concat()
                               - number()
- contains()
                               - position()
                               - round()
- count()
- false()
                               - starts-with()
- floor()
                               - string()
                               - string-length()
- id()
                               - substring()
- lang()
                               - substring-after()
- last()
- local-name()
                               - substring-before()
- name()
                               - sum()
- namespace-uri()
                               - translate()
                               - true()
```

Biblioteka standardnih funkcija

primeri

- position() vraća redni broj čvora u tekućem kontekstu
 - //book[position()=last()] poslednji book element
- count() vraća broj elemenata u skupu čvorova
 - count(//book) broj book elemenata
- substring() vraća podstring deo tekstualnog sadržaja
 - substring('trt mrt', 1, 3) vraća trt

Označavanje više putanja odjednom

- navođenjem više XPath izraza povezanih operatorom |
- rezultat je unija skupova čvorova dobijenih osnovnim izrazima
- primeri
 - //book/title | //book/price
 - svi title elementi deca book-a i svi price elementi deca book-a
 - -//title | //price
 - svi *title* elementi i svi *price* elementi
 - /bookstore/book/title | //price
 - svi title elementi book elementa bookstore korena i svi price elementi