

IT452 Advanced Web and Internet Systems

Set 7: XML and XPath (Chapter 14 of text)

Example XML

```
<?xml version="1.0" encoding="UTF-8"?>
<RDF>
  <channel about="http://web2.0thebook.org/channel.rss">
    <title>Planet web2.0thebook</title>
    <title>This is our alternate title </title>
    <link>http://web2.0thebook.org/</link>
    <description>Aggregated content relevant to the upcoming book "Professional Web 2.0
      Programming".</description>
  </channel>
  <item about="http://www.orbeon.com/blog/2006/06/13/firebug-a-must-have-firefox-extension-for-web-
    developers/">
    <title>XForms Everywhere Â» FireBug: A Must-Have Firefox Extension for Web Developers</title>
    <link>http://www.orbeon.com/blog/2006/06/13/firebug-a-must-have-firefox-extension-for-web-
      developers/</link>
    <description>Alessandro Vernet recommends FireBug, à200\234an absolute godsendà200\235, the
      à200\234reatest web
      developer extension out thereà200\235, an à200\234awesomeà200\235,
      à\200\234phenomenalà200\235, and à200\234absolutely, completely
      brilliantà200\235 extension.</description>
  </item>
  <item about="http://eric.van-der-vlist.com/blog/2504_Web_2.0_at_XML_Prague.item">
    <title>Web 2.0 at Prague</title>
    <link>http://eric.van-der-vlist.com/blog/2504_Web_2.0_at_XML_Prague.item</link>
    <description>Eric van der Vlist will do a presentation about Web 2.0 at XML Prague
      2006.</description>
  </item>
  <item about="http://www.orbeon.com/blog/2006/06/10/unicode-in-java-not-so-fast/">
    <title>XForms Everywhere Â» Unicode in Java: not so fast (but XML is better)!</title>
    <link>http://www.orbeon.com/blog/2006/06/10/unicode-in-java-not-so-fast/</link>
  </item>
</RDF>
```

XML with namespaces

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://realPrefixRemovedForThisDemo.org#" xmlns="http://purl.org/rss/1.0/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <channel rdf:about="http://web2.0thebook.org/channel.rss">
    <title>Planet web2.0thebook</title>
    <link>http://web2.0thebook.org/</link>
    <description>Aggregated content relevant to the upcoming book "Professional Web 2.0
      Programming".</description>
  </channel>
  <item>
    rdf:about="http://www.orbeon.com/blog/2006/06/13/firebug-a-must-have-firefox-extension-for-web-developers/>
    <title>XForms Everywhere A» FireBug: A Must-Have Firefox Extension for Web Developers</title>
    <link>http://www.orbeon.com/blog/2006/06/13/firebug-a-must-have-firefox-extension-for-web-developers/</link>
    <description>Alessandro Vernet recommends FireBug, à\200\234an absolute godsendà\200\235, the
      à\200\234greatest web
      developer extension out thereà\200\235, an à\200\234awesomeà\200\235, à\200\234phenomenalà\200\235, and
      à\200\234absolutely, completely
      brilliantà\200\235 extension.</description>
    <dc:creator>evlist</dc:creator>
    <dc:date>2006-06-15T05:56:16Z</dc:date>
    <dc:subject>javascript tools web2.0thebook webdev</dc:subject>
  </item>
  <item rdf:about="http://eric.van-der-vlist.com/blog/2504_Web_2.0_at_XML_Prague.item">
    <title>Web 2.0 at Prague</title>
    <link>http://eric.van-der-vlist.com/blog/2504_Web_2.0_at_XML_Prague.item</link>
    <description>Eric van der Vlist will do a presentation about Web 2.0 at XML Prague 2006.</description> ...
  
```

Which mean the same thing?

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://realPrefixRemovedForThisDemo.org#" xmlns="http://purl.org/rss/1.0/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <channel rdf:about="http://web2.0thebook.org/channel.rss">
    <title>Planet web2.0thebook</title>
    <link>http://web2.0thebook.org/</link>
  </channel> ...

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://realPrefixRemovedForThisDemo.org#" xmlns:rss="http://purl.org/rss/1.0/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rss:channel rdf:about="http://web2.0thebook.org/channel.rss">
    <rss:title>Planet web2.0thebook</rss:title>
    <rss:link>http://web2.0thebook.org/</rss:link>
  </rss:channel> ...

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://realPrefixRemovedForThisDemo.org#" xmlns:rss="http://blahblahblah.com/stuff"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rss:channel rdf:about="http://web2.0thebook.org/channel.rss">
    <rss:title>Planet web2.0thebook</rss:title>
    <rss:link>http://web2.0thebook.org/</rss:link>
  </rss:channel> ...

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://realPrefixRemovedForThisDemo.org#" xmlns:dog="http://purl.org/rss/1.0/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <dog:channel rdf:about="http://web2.0thebook.org/channel.rss">
    <dog:title>Planet web2.0thebook</dog:title>
    <dog:link>http://web2.0thebook.org/</dog:link>
  </dog:channel> ...

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://realPrefixRemovedForThisDemo.org#" xmlns:dog="http://blahblahblah.com/stuff"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <dog:channel rdf:about="http://web2.0thebook.org/channel.rss">
    <dog:title>Planet web2.0thebook</dog:title>
    <dog:link>http://web2.0thebook.org/</dog:link>
  </dog:channel> ...
```

XPath

```
/  
/RDF  
/RDF/channel/title  
channel/title  
channel//title  
item[@title='Jaw surgery']
```

More complex:

```
/rdf:RDF/rss:item[@rdf:about=current()]/@rdf:resource]
```

Exercise: XPath (part 1)

```
<booklist listtitle="Science Fiction">  
  <book>  
    <title>The Naked Sun</title>  
    <author>Isaac Asimov</author>  
    <isbn>0553293397</isbn>  
    <price>30</price> <!-- add by hand to online demo -->  
  </book>  
  
  <book>  
    <title>Foundation's Triumph</title>  
    <author>David Brin</author>  
    <isbn>0061056391</isbn>  
    <price>20</price> <!-- add by hand to online demo -->  
  </book>  
  
  <book>  
    <title>Snow Crash</title>  
    <author>Neal Stephenson</author>  
    <isbn>0553380958</isbn>  
  </book>  
</booklist>
```

Demo from: <http://www.futurelab.ch/xmlkurs/xpath.en.html>

Exercise: XPath (part 2)

Write XPath to find:

1. All the book elements
2. The ISBN of all the books
3. Book elements with author Isaac Asimov
4. The title of all books priced > 20