XML
&
Allied Technologies

XPath

What is XPath?

- XPath stands for XML Path Language.
- XPath uses "path-like" syntax to identify and navigate nodes in an XML document.
- XPath is a W3C recommendation.

What is XPath?(cont.)

- XPath uses path expressions to select nodes or node-sets in an XML document.
- These path expressions look very much like the path expressions you use with traditional computer file systems:

XPath Standard Functions

- XPath includes over 200 built-in functions.
- There are functions for string values, numeric values, booleans, date and time comparison, node manipulation, sequence manipulation, and much more.
- Today XPath expressions can also be used in JavaScript, Java, PHP, Python, C and C++, and lots of other languages.

XPath is a W3C Recommendation

- XPath 1.0 became a W3C Recommendation on November 16, 1999.
- XPath 2.0 became a W3C Recommendation on January 23, 2007.
- XPath 3.0 became a W3C Recommendation on April 8, 2014.

XPath Nodes

- In XPath, there are **seven** kinds of nodes:
  - Element.
  - Attribute.
  - Text (The text content of a node).
  - Namespace (xmlns="namespace").
  - Processing-instruction.
  - Comment.
  - Document nodes.
- XML documents are treated as trees of nodes. The topmost element of the tree is called the root element.

XPath Nodes (cont.)

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

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

root element node

Attribute node

Element node

Relationship of Nodes

<book> (Parent)

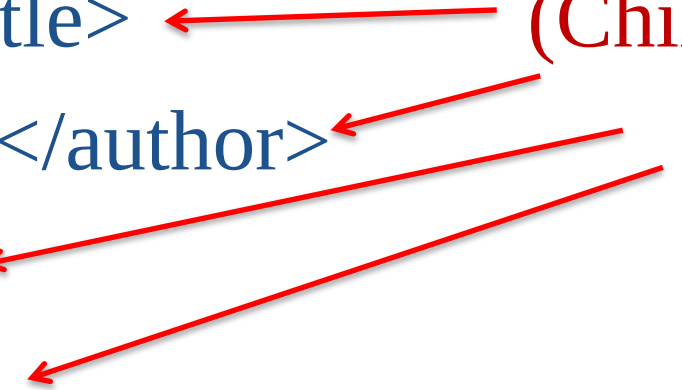
<title>Harry Potter</title> (Children)

<author>J K. Rowling</author>

<year>2005</year>

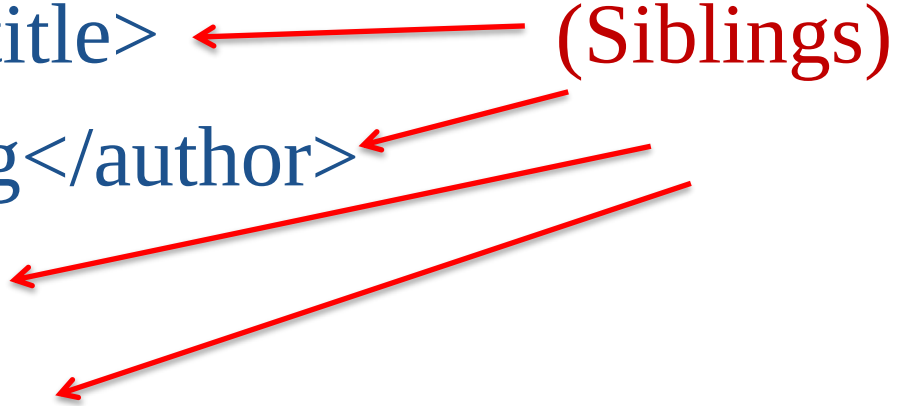
<price>29.99</price>

</book>

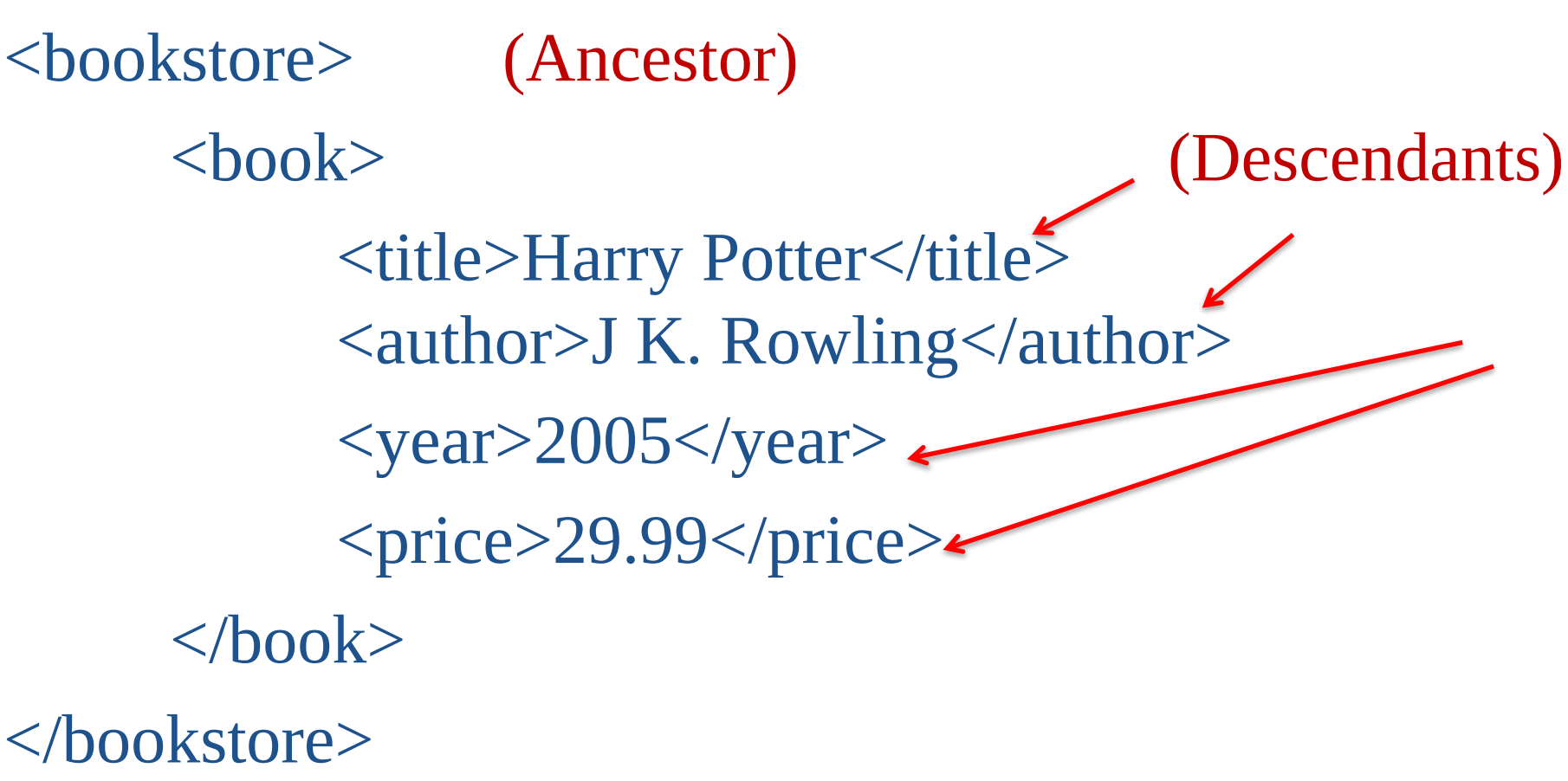


Relationship of Nodes (cont.)

```
<book>
  <title>Harry Potter</title>
  <author>J K. Rowling</author>
  <year>2005</year>
  <price>29.99</price>
</book>
```



Relationship of Nodes (cont.)



XPath Expressions

Upon being processed and evaluated, XPath expressions result in a data object of one of the following types:

- Node set: A collection of nodes
- String: A text string
- Boolean: A true/false value
- Number: A floating-point number

Location Path Expression

A location path can be absolute or relative.

An absolute location path starts with a slash (/) and a relative location path does not.

In both cases the location path consists of one or more steps, each separated by a slash:

```
An absolute location path:

/step/step/...

A relative location path:

step/step/...
```

XPath Syntax

Expression	Description
nodename	Selects all nodes with the name "*nodename*"
/	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

XPath Syntax (cont.)

```
<?xml version="1.0" encoding="UTF-8"?>
<bookstore>
<book>
  <title lang="en">Harry Potter</title>
  <price>29.99</price>
</book>
<book>
  <title lang="en">Learning XML</title>
  <price>39.95</price>
</book>
</bookstore>
```

XPath Syntax (cont.)

Path Expression	Result
bookstore	Selects all nodes with the name "bookstore"
/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

XPath Operators

Operator	Description	Example
|	Computes two node-sets	//book | //cd
+	Addition	6 + 4
-	Subtraction	6 - 4
*	Multiplication	6 * 4
div	Division	8 div 4
=	Equal	price=9.80
!=	Not equal	price!=9.80
<	Less than	price<9.80
<=	Less than or equal to	price<=9.80
>	Greater than	price>9.80
>=	Greater than or equal to	price>=9.80
or	or	price=9.80 or price=9.70
and	and	price>9.00 and price<9.90
mod	Modulus (division remainder)	5 mod 2

XPath Syntax (cont.)

Path Expression	Result
/bookstore/book[1]	Selects the first book element that is the child of the bookstore element. **Note:** In IE 5,6,7,8,9 first node is[0], but according to W3C, it is [1]. To solve this problem in IE, set the SelectionLanguage to XPath: *In JavaScript: xml*.setProperty("SelectionLanguage","XPath");
/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='en']	Selects all the title elements that have a "lang" attribute with a value of "en"
/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

XPath Examples

•Using the books.xml sample file.

1. Select all the titles:
   - /bookstore/book/title.
2. Select the title of the first book:
   - /bookstore/book[1]/title
3. Select all the prices:
   - /bookstore/book/price/text().
4. Select price nodes with price>35:
   - /bookstore/book[price>35]/price.

Thanks