

ICS313 :5.0



Web Programming

Lecturer: Toby Daniel

Introduction to XML

Extensible Markup Language

- XML is a markup language similar to HTML and is used to describe data.

HTML and XML

- XML does not replace HTML
 - (it is not a *better* version)
- The goal of HTML is to display data.
- The goal of XML is to describe data.
- XML uses tags in a similar way to HTML.
- XML is plain text.
- XML tags are not predefined in XML. You must define your own tags
 - (therefore it is extensible!)

XML example

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

```
<message date="24/01/08">
```

```
<to>ics313 students</to>
```

```
<from>Toby</from>
```

```
<heading>Exam Schedule</heading>
```

```
<body>Log into E-registrar to see the new exam  
schedule.</body>
```

```
</message>
```

Separating Data from Display from Style

- HTML pages are used to display data.
- CSS is used to add style to the displayed data.
- Data is often stored inside HTML pages.
- With XML this data can now be stored in a separate XML file.
- XML data can also be stored inside HTML pages as "Data Islands".

Data Exchange

- Computer systems and databases usually contain data in incompatible formats.
- This leads to difficulties when trying to exchange data between systems over the Internet.
- Converting the data to XML can reduce this problem and create data that can be read by different types of applications.

Data Storage

- XML can also be used to store data in files or in databases.
- Applications can be written to store and retrieve information from the store and then correctly display the data.

XML Rules

- In XML all elements must have a closing tag.
 - `<message>blah blah</message>`
- XML tags are case sensitive.
 - `<message>` is different from `<Message>`
- In XML all elements must be properly nested
 - `<message><time>08:00</message></time>` ❌
 - `<message><time>08:00</time></message>` ✅

XML Format

- XML must contain a single tag pair to define the root element.
- All other elements must be nested within the root element.
- All elements can have sub (child) elements.

```
<root>  
  <child>  
    <subchild>  
    </subchild>  
  </child>  
</root>
```

XML Examples

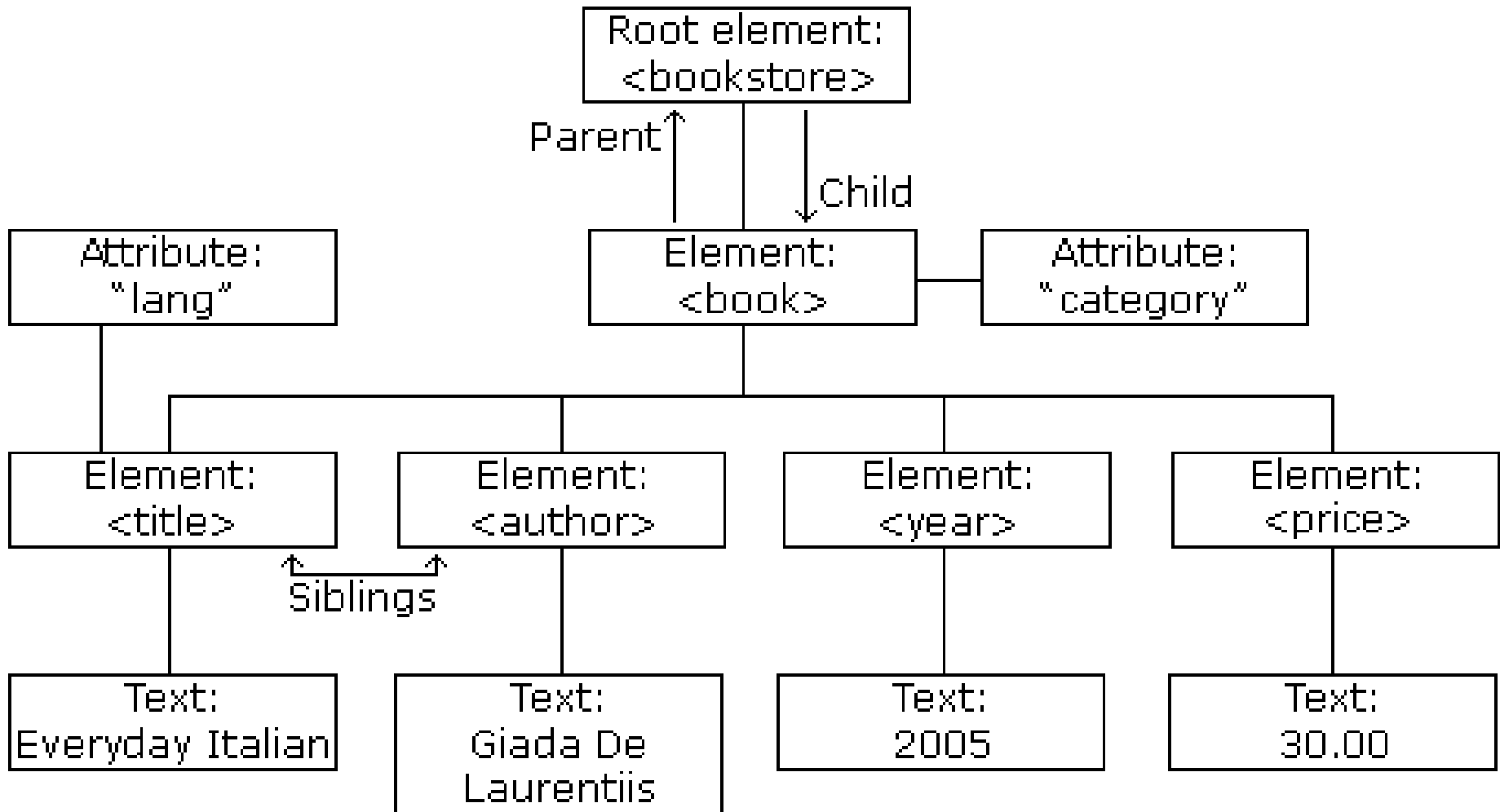
- The next slide shows a small example of what could be a larger XML file from a book shop.
- Notice the root and child elements.

```
<?xml version="1.0"?>
<bookstore>
  <book category="Computer">
    <title lang="en">Distributed Systems</title>
    <author>Andrew Tanenbaum</author>
    <author>Maarten Van Steen</author>
    <year>2002</year>
    <price>90.00</price>
  </book>
  <book category="Computer">
    <title lang="en">PHP Functions</title>
    <author>Zak Greant</author>
    <year>2002</year>
    <price>49.99</price>
  </book>
</bookstore>
```

DOM

- DOM = Document Object Model
- The XML Document Object Model (XML DOM) defines a standard way for accessing and manipulating XML documents.
- The DOM presents an XML document as a tree-structure (a node tree), with the elements, attributes, and text defined as nodes.

DOM for the bookstore example



XML DOM Text Nodes

- Text is always stored in **text nodes**.
- A common error in looking at the DOM is to navigate to an element node and expect it to contain the text.
- Remember, even the simplest element node has a text node under it.
- For example:
 - `<year>2005</year>`
 - there is an element node [year]
 - and a text node under it with the text [2005]

Why do we need the DOM?

- You need the DOM to:
 - find Nodes
and
 - access Nodes
- .. in an XML document.

XML and your Web Browser

- If you have a modern web browser then it will read XML files.
- It will display the document tree with child nodes indented from the left margin.
- It will also show any syntax errors in those files.

Exercise

- Take the bookstore example (slide 11), cut and paste it into your text editor and save it with .xml extension. Then open it in your web browser.

Reasons to use XML

XML is Used to Create New Internet Languages

- RSS languages for news feeds
- WSDL for describing available web services
- WAP and WML as markup languages for handheld devices
- RDF and OWL for describing resources and ontology
- SMIL for describing multimedia for the web

Reasons to use XML

- XML is independent of hardware, software and applications.
- XML can make your data more available and useful – sharing, transferring, merging
- XML helps with platform changes. Data doesn't have to be converted for it to maintain it's meaning.