*e***X***tensible* **M***arkup* **L***anguage* (XML)

Original slides by:

Albert Beng Kiat Tan

Ayzer Mungan

Edwin Hendriadi

What is XML?

* *eXtensible Markup Language*
* Markup language for documents containing semi-structured information
* Ensures portability of information
  * Both human and machine readable
* Roughly, the "alphabet" of web service interactions

XML….

- Based on Standard Generalized Markup Language (SGML)
- Version 1.0 introduced by World Wide Web Consortium (W3C) in 1998
- Bridge for data exchange on

  the Web

SGML

HTML   XML
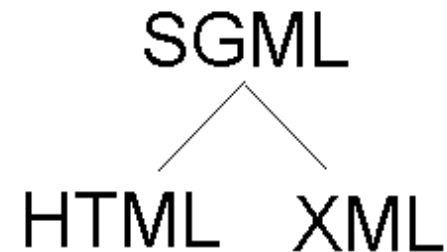
Comparison

XML

* Unspecified set of tags

* Content oriented

HTML

* Fixed set of tags

* Presentation oriented

Authoring XML Elements

* An XML element is made up of a start tag, an end tag, and data in between.
* Example:

    <director> Matthew Dunn </director>
* Example of another element with the same value:

    <actor> Matthew Dunn </actor>
* XML tags are case-sensitive:

    <CITY> <City> <city>
* XML can abbreviate empty elements, for example:
can be abbreviated to

Authoring XML Elements (cont'd)

* An attribute is a name-value pair separated by an equal sign (=).

* Example:

    <City ZIP="94608"> Emeryville </City>

* Attributes are used to attach additional, secondary information to an element.

Authoring XML Documents

* A basic XML document is an XML element that can, but might not, include nested XML elements, preceded by a prologue

* Example:

```
<?xml version="1.0" encoding="UTF-8"?>
<books>
    <book isbn="123">
        <title> Second Chance </title>
        <author> Matthew Dunn </author>
    </book>
</books>
```

Authoring XML Documents (cont'd)

- Authoring guidelines:
  - All elements must have an end tag.
  - All elements must be cleanly nested (overlapping elements are not allowed).
  - All attribute values must be enclosed in quotation marks.
  - Each document must have a unique first element, the root node.

Comments

- In XML, comments are as follows
  - <!-- comment -->
  - A comment should not contain --
  - Comments are not allowed inside other markups

Namespaces

* Designers can choose the names of their tags
* Name clashes may happen
* Namespaces used to avoid ambiguities
* Namespaces are URIs
  * It is not needed that they point to actual resources

Namespaces: example

* ```xml
  <?xml version="1.0" encoding="UTF-8"?>
  <books xmlns="http://www.mylibrary.com" >
      <book isbn="123">
          <title xmlns="http://www.mytitles.com">
          Second Chance </title>
          <author> Matthew Dunn </author>
      </book>
  </books>
  ```

* We use the xmlns attribute

* The scope is the element where it is declared and its sub elements

Namespaces: shortcut

- One can define a shortcut for a namespace and then use it

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

  <bo:books xmlns:bo="http://www.mylibrary.com"
            xmlns:ti="http://www.mytitles.com">
      <bo:book isbn="123">
          <ti:title> Second Chance </ti:title>
          <bo:author> Matthew Dunn </bo:author>
      </bo:book>
  </bo:books>