

XML

Session 3: Introduction to Document Type Definition (DTD)

What is a DTD

- Document Type Definition
- Sets the rules for structure and type of content for an xml document, ensuring “validity”
- Has its own syntax – but is not XML
 - Syntax derived from Standard Generalized Markup Language (SGML)

Initial XML document

```
<?xml version="1.0" encoding="UTF-8"?>
<photos>
  <photoitem ref="logo" folder="images" type="jpg">
    <title>Logo</title>
    <text>
      <p>My company logo image</p>
    </text>
    <captions>
      <caption>Company Logo</caption>
    </captions>
  </photoitem>
</photos>
```

Filename: Photos.xml

DTD that validates the XML doc

```
<?xml version="1.0" encoding="UTF-8"?>
<!ELEMENT photos (photoitem*)>
<!ELEMENT photoitem (title, text?, captions)>
<!ATTLIST photoitem
  ref ID #REQUIRED
  folder NMTOKEN #REQUIRED
  type (jpg | gif | png) "jpg"
>
<!ELEMENT title (#PCDATA)>
<!ELEMENT text (p+ | ul*)*>
<!ELEMENT p (#PCDATA)*>
<!ELEMENT captions (caption+)>
<!ELEMENT caption (#PCDATA)>
<!ELEMENT ul (#PCDATA)>
```

Filename: Photos.dtd

** Note: NMTOKEN stands for Name Token and indicates that the value of the element may NOT have spaces in it and must conform to XML elements naming conventions (i.e. elements can only start with a letter, an underscore or a colon).*

XML document to create a DTD from

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE photos SYSTEM "Photos.dtd">
<photos>
  <photoitem ref="logo" folder="images" type="jpg">
    <title>Logo</title>
    <text>
      <p>My company logo image</p>
    </text>
    <captions>
      <caption>Company Logo</caption>
    </captions>
  </photoitem>
</photos>
```

DTD Syntax – Elements

- To specify an element:
 - `<!ELEMENT [element_name] [child_definitions]>`
- An element name must be a valid XML name
- The child definitions can be as follows:
 - #PCDATA (Parsed character data)
Example: `<!ELEMENT title (#PCDATA)>`
 - Child Elements
Example: `<!ELEMENT captions (caption)>`
 - Sequences
Example: `<!ELEMENT photoitem (title, text?, captions)>`

- Number of children
 - ? Zero or one of the element
 - * Zero or more of the element
 - + One or more of the element

Example DTD:

```
<?xml version="1.0" encoding="UTF-8"?>
<!ELEMENT customerorder (companyname, mailingaddress,
    billingaddress?, orderitem+)>
<!ELEMENT mailingaddress (addressline+, direction*)>
<!ELEMENT billingaddress (addressline)>
<!ELEMENT companyname (#PCDATA)>
<!ELEMENT addressline (#PCDATA)>
<!ELEMENT direction (#PCDATA)>
<!ELEMENT orderitem (#PCDATA)>
```

Filename: CustomerOrders.dtd

Sample XML document

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE customerorder SYSTEM "CustomerOrders.dtd">
<customerorder>
  <companyname>Despatch International Ltd</companyname>
  <mailingaddress>
    <addressline>1265 Manor Street</addressline>
    <addressline>Berkley, CA 94710</addressline>
    <direction>Take right turning at Hollow Way</
direction>
    <direction>Turn left at notting hill junction</
direction>
  </mailingaddress>
  <orderitem>XML How To Program Deitel</orderitem>
  <orderitem>CD Album Blue 2004</orderitem>
</customerorder>
```

Filename: CustomerOrders.xml

DTD Syntax – Elements

- Choices and grouping (parentheses)
`<!ELEMENT employee (firstName, lastName, (employeeid | nationalinsurance))>`

- Empty Elements
`<!ELEMENT img EMPTY>`

XML examples:

``

or

``

- Mixed content
 - DTD

```
<?xml version="1.0" encoding="UTF-8"?>  
<!ELEMENT summarymessage (#PCDATA | warningmessage)*>  
<!ELEMENT warningmessage (#PCDATA)>
```

Filename: WarningMessages.dtd

- Sample XML document:

```
<?xml version="1.0" encoding="UTF-8"?>  
<!DOCTYPE summarymessage SYSTEM "WarningMessages.dtd">  
<summarymessage>You have entered invalid values in the  
fields below:  
  <warningmessage>Start date must be greater than 31-Jan-  
2006</warningmessage>  
  <warningmessage>End date must be less or equal to 31-  
Jan-2007</warningmessage>  
</summarymessage>
```

Filename: WarningMessages.xml

DTD Syntax – Attributes

- To specify an attribute for an element, you must first define the element, then:
 - DTD Syntax: `<!ATTLIST [element_name] [attribute_list]>`
Example DTD: `<!ATTLIST client id CDATA #REQUIRED>`
Sample XML: `<client id="NI1234"></client>`
- The attribute list is:
 - `[attribute_name] [attribute_type] [attribute_default]`
 - An attribute name must be a valid XML name
 - An attribute type can be:
 - CDATA – any string of text
 - NMTOKEN – a name token
 - NMTOKENS – sequence of name tokens separated by spaces
 - ID – a unique reference for the element in the document

DTD Syntax – Attributes

- IDREF – a reference to ID
- IDREFS – sequence of IDREF separated by spaces
- Enumeration

Example DTD:

```
<!ATTLIST date month (Jan|Feb|Mar|Apr|May|Jun|Jul|Aug|Sep|Oct|Nov|Dec)  
#REQUIRED>
```

Sample XML:

```
<date month="Jan"></date>
```

– Attribute defaults can be:

- #IMPLIED – the attribute is optional
- #REQUIRED – the attribute must be set
- #FIXED – the attribute can only have the specified value
- Default value specified in quotes

Example DTD:

```
<!ATTLIST img width CDATA "40">
```

Sample XML:

```
<img width="40"></img>
```