W3C XML Schema Documents

Defining an Element in XML Schema

- In XML Schema, the `element` tag defines an element to be included in an XML document that conforms to the schema.
- An element's `type` attribute indicates the data type that the element may contain.
- Possible types include XML Schema–defined types and user-defined types.

W3C XML Schema Documents (Cont.)

- Two categories of types exist in XML Schema—**simple types** and **complex types**.
- They differ only in that simple types cannot contain attributes or child elements and complex types can.
- A user-defined type that contains attributes or child elements must be defined as a complex type.
- The `sequence` allows you to specify the sequential order in which child elements must appear.
- Attribute `minOccurs` specifies the minimum occurrences of an element. `maxOccurs` specifies the maximum; it can be set to `unbounded`.
- Both of these attributes have default values of `1`.

W3C XML Schema Documents (Cont.)

A Closer Look at Types in XML Schema

- Every element in XML Schema has a type that is either built-in or user-defined.
- Every simple type defines a **restriction** on a previously defined type.
- Complex types are divided into two groups—those with **simple content** and those with **complex content**.
  - Both can contain attributes.
  - Only complex content can contain child elements.
- Complex types with simple content must extend or restrict some other existing type.
- Complex types with complex content do not have this limitation.

Outline

- The schema document in Fig. 20.6 creates both simple types and complex types.

computer.xsd

(1 of 2)

```
1  <?xml version = "1.0"?>
2  <!-- Fig. 20.6: computer.xsd -->
3  <!-- W3C XML Schema document -->
4
5  <schema xmlns = "http://www.w3.org/2001/XMLSchema"
6     xmlns:computer = "http://www.deitel.com/computer"
7     targetNamespace = "http://www.deitel.com/computer">
8
9     <simpleType name = "gigahertz">
10       <restriction base = "decimal">
11          <minInclusive value = "2.1"/>
12       </restriction>
13    </simpleType>
14
15    <complexType name = "CPU">
16       <simpleContent>
17          <extension base = "string">
```

Specify the base type as `decimal`.

Restrict the value to be at least `2.1`.

Create a simple type using the `simpleType` element.

Declare a `complexType` named `CPU` that has `simpleContent` and a base type of `string`.

The `extension` element with attribute `base` sets the base type.

Fig. 20.6 | XML Schema document defining simple and complex types. (Part 1 of 2.)

Outline

computer.xsd (2 of 2)

```
18            <attribute name = "model" type = "string"/>
19         </extension>
20      </simpleContent>
21   </complexType>
22
23   <complexType name = "portable">
24      <all>
25         <element name = "processor" type = "computer:CPU"/>
26         <element name = "monitor" type = "int"/>
27         <element name = "CPUSpeed" type = "computer:gigahertz"/>
28         <element name = "RAM" type = "int"/>
29      </all>
30      <attribute name = "manufacturer" type = "string"/>
31   </complexType>
32
33   <element name = "laptop" type = "computer:portable"/>
34 </schema>
```

The `attribute` element specifies that an element of type `CPU` and may contain a `model` attribute that is of type `string`.

The element `all` encloses elements that must each be included once in the document. These elements can be included in any order.

`portable` contains an attribute of type `string` named `manufacturer`.

Declare the actual element, called `laptop`, that uses the three types defined in the schema.

Fig. 20.6 | XML Schema document defining simple and complex types. (Part 2 of 2.)

XML Schema Documents (Cont.)

- A document that conforms to a schema is known as an **XML instance document**.
- Create a simple type using the `simpleType` element.
- Simple types are restrictions of a type typically called a **base type**.
- The `extension` element with attribute `base` sets the base type.
- The element `all` encloses elements that must each be included once in the document. These elements can be included in any order.

Outline

- Figure 20.7 uses the `laptop` element defined in the `computer.xsd` schema.

`laptop.xml`

```
1  <?xml version = "1.0"?>
2  <!-- Fig. 20.7: laptop.xml              -->
3  <!-- Laptop components marked up as XML -->
4
5  <computer:laptop xmlns:computer = "http://www.deitel.com/computer"
6     manufacturer = "IBM">
7
8     <processor model = "Centrino">Intel</processor>
9     <monitor>17</monitor>
10    <CPUSpeed>2.4</CPUSpeed>
11    <RAM>256</RAM>
12 </computer:laptop>
```

Fig. 20.7 | XML document using the laptop element defined in `computer.xsd`.

W3C XML Schema Documents (Cont.)

Automatically Creating Schemas using Visual Studio

- Visual Studio includes a tool that allows you to create a schema from an existing XML document, using the document as a template.
- With an XML document open, select **XML > Create Schema** to use this feature.

Extensible Stylesheet Language and XSL Transformations

- **Extensible Stylesheet Language** (**XSL**) documents specify how programs are to render XML document data.
- XSL is a group of three technologies—**XSL-FO** (**XSL Formatting Objects**), **XPath** (**XML Path Language**) and **XSLT** (**XSL Transformations**).
  - XSL-FO is a vocabulary for specifying formatting.
  - XPath is a string-based language used by XML to locate structures and data in XML documents.
  - XSL Transformations (XSLT) is a technology for transforming XML documents into other documents.
- XSLT allows you to convert an XML document to an **XHTML** (**Extensible HyperText Markup Language**) document for display in a web browser.

Extensible Stylesheet Language and XSL Transformations (Cont.)

- Transforming an XML document using XSLT involves two tree structures—the **source tree** and the **result tree**.
- XPath is used to locate parts of the source-tree document that match **templates** defined in an **XSL style sheet**.
- When a match occurs, the matching template executes and adds its result to the result tree.
- The XSLT does not analyze every node of the source tree; it selectively navigates the source tree using XSLT's `select` and `match` attributes.
- For XSLT to function, the source tree must be properly structured.

A Simple XSL Example

- Figure 20.8 lists an XML document that describes various sports.

`sports.xsl`

(1 of 2)

```
1  <?xml version = "1.0"?>
2  <?xml-stylesheet type = "text/xsl" href = "sports.xsl"?>
3
4  <!-- Fig. 20.8: sports.xml -->
5  <!-- Sports Database -->
6
7  <sports>
8     <game id = "783">
9        <name>Cricket</name>
10
11       <paragraph>
12          More popular among Commonwealth nations.
13       </paragraph>
14    </game>
15
16    <game id = "239">
17       <name>Baseball</name>
18
```

This **processing instruction** (**PI**) specifies the location of the XSL style sheet `sports.xsl`, which will be used to transform the XML document.

Fig. 20.8 | XML document that describes various sports. (Part 1 of 2.)

```
19       <paragraph>
20          More popular in America.
21       </paragraph>
22    </game>
23
24    <game id = "418">
25       <name>Soccer (Futbol)</name>
26
27       <paragraph>
28          Most popular sport in the world.
29       </paragraph>
30    </game>
31 </sports>
```

sports.xsl

(2 of 2)

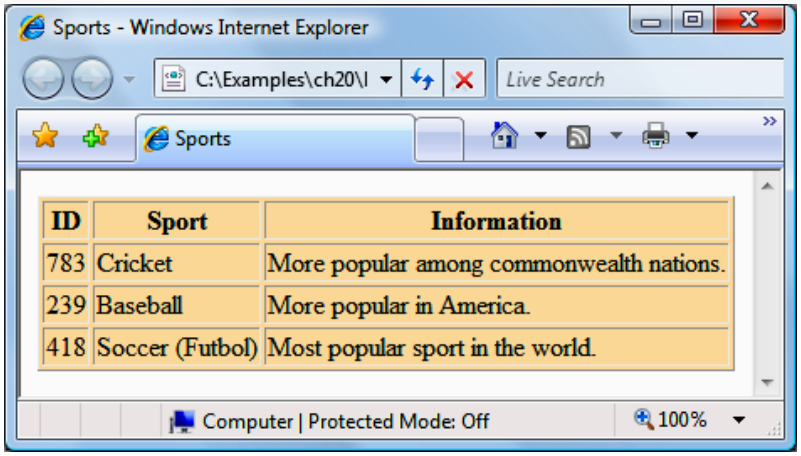

Fig. 20.8 | XML document that describes various sports. (Part 2 of 2.)

- Figure 20.9 shows the XSL document for transforming the structured data of the XML document of Fig. 20.8 into an XHTML document for presentation.

Outline

sports.xsl

(1 of 2)

```
1  <?xml version = "1.0"?>
2  <!-- Fig. 20.9: sports.xsl -->
3  <!-- A simple XSLT transformation -->
4
5  <!-- reference XSL style sheet URI -->
6  <xsl:stylesheet version = "1.0"
7     xmlns:xsl = "http://www.w3.org/1999/XSL/Transform">
8
9     <xsl:output method = "xml" omit-xml-declaration = "no"
10       doctype-system =
11          "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd"
12       doctype-public = "-//W3C//DTD XHTML 1.0 Strict//EN"/>
13
14    <xsl:template match = "/"> <!-- match root element -->
15
16    <html xmlns = "http://www.w3.org/1999/xhtml">
17       <head>
18          <title>Sports</title>
19       </head>
20
```

The style sheet begins with the `stylesheet` start tag. Attribute `version` specifies the XSLT version.

Attribute `method` is assigned `"xml"`, which indicates that XML is being output to the result tree. Attribute `omit-xml-declaration` specifies that we should not want to omit the XML declaration

Use element `xsl:output` to write an XHTML document type declaration (`DOCTYPE`) to the result tree.

Attributes `doctype-system` and `doctype-public` write the `DOCTYPE` DTD information to the result tree.

Use the `match` attribute to select the **document root** of the XML source document.

The XSLT processor writes this XHTML to the result tree exactly as it appears in the XSL document.

Fig. 20.9 | XSLT that creates elements and attributes in an XHTML document. (Part 1 of 2.)

Outline

sports.xsl

(2 of 2)

```
21       <body>
22          <table border = "1" style = "background-color: wheat">
23             <thead>
24                <tr>
25                   <th>ID</th>
26                   <th>Sport</th>
27                   <th>Information</th>
28                </tr>
29             </thead>
30
31             <!-- insert each name and paragraph element value -->
32             <!-- into a table row. -->
33             <xsl:for-each select = "/sports/game">
34                <tr>
35                   <td><xsl:value-of select = "@id"/></td>
36                   <td><xsl:value-of select = "name"/></td>
37                   <td><xsl:value-of select = "paragraph"/></td>
38                </tr>
39             </xsl:for-each>
40          </table>
41       </body>
42    </html>
43
44    </xsl:template>
45 </xsl:stylesheet>
```

Use element `value-of` to retrieve an attribute's value using the XPath symbol `@`.

Element `xsl:for-each` iterates through the source XML document, searching for the element specified by the `select` attribute.

Fig. 20.9 | XSLT that creates elements and attributes in an XHTML document. (Part 2 of 2.)

Extensible Stylesheet Language and XSL Transformations (Cont.)

- The style sheet begins with the `stylesheet` start tag. Attribute `version` specifies the XSLT version.
- The `DOCTYPE` identifies XHTML as the type of the resulting document.
- Attributes `doctype-system` and `doctype-public` write the `DOCTYPE` DTD information to the result tree.
- XSLT uses **templates** (i.e., `xsl:template` elements) to describe how to transform particular nodes from the source tree to the result tree.
- A template is applied to nodes that are specified in the `match` attribute.

Extensible Stylesheet Language and XSL Transformations (Cont.)

- The XPath character `/` (a forward slash) is used as a separator between element names.
- In XPath, a leading forward slash specifies that we are using **absolute addressing** (i.e., we are starting from the root).
- Element `xsl:for-each` iterates through the source XML document, searching for the element specified by the `select` attribute.
- Use element `value-of` to retrieve an attribute's value using the XPath symbol `@`.
- When an XPath expression has no beginning forward slash, the expression uses **relative addressing**.

Using XSLT to Sort and Format Data

- Figure 20.10 presents an XML document (`sorting.xml`) that marks up information about a book.

`sorting.xml`

(1 of 2)

```
1  <?xml version = "1.0"?>
2  <!-- Fig. 20.10: sorting.xml -->
3  <!-- XML document containing book information -->
4
5  <?xml-stylesheet type = "text/xsl" href = "sorting.xsl"?>
6
7  <book isbn = "999-99999-9-X">
8     <title>Deitel&apos;s XML Primer</title>
9
10    <author>
11       <firstName>Jane</firstName>
12       <lastName>Blue</lastName>
13    </author>
14
15    <chapters>
```

Fig. 20.10 | XML document containing book information. (Part 1 of 2.)

sorting.xml

(2 of 2)

```
16       <frontMatter>
17          <preface pages = "2" />
18          <contents pages = "5" />
19          <illustrations pages = "4" />
20       </frontMatter>
21
22       <chapter number = "3" pages = "44">Advanced XML</chapter>
23       <chapter number = "2" pages = "35">Intermediate XML</chapter>
24       <appendix number = "B" pages = "26">Parsers and Tools</appendix>
25       <appendix number = "A" pages = "7">Entities</appendix>
26       <chapter number = "1" pages = "28">XML Fundamentals</chapter>
27    </chapters>
28
29    <media type = "CD" />
30 </book>
```

Fig. 20.10 | XML document containing book information. (Part 2 of 2.)

- XSL style sheet can sort an XML file's data for presentation purposes.

Outline

- Figure 20.11 presents an XSL document (`sorting.xsl`) for transforming `sorting.xml` (Fig. 20.10) to XHTML.

`sorting.xsl`

(1 of 5)

```
1  <?xml version = "1.0"?>
2  <!-- Fig. 20.11: sorting.xsl -->
3  <!-- Transformation of book information into XHTML -->
4
5  <xsl:stylesheet version = "1.0" xmlns = "http://www.w3.org/1999/xhtml"
6     xmlns:xsl = "http://www.w3.org/1999/XSL/Transform">
7
8     <!-- write XML declaration and DOCTYPE DTD information -->
9     <xsl:output method = "xml" omit-xml-declaration = "no"
10       doctype-system = "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd"
11       doctype-public = "-//W3C//DTD XHTML 1.1//EN"/>
12
13    <!-- match document root -->
14    <xsl:template match = "/">
15       <html>
16          <xsl:apply-templates/>
17       </html>
18    </xsl:template>
19
```

The `<xsl:apply-templates/>` element specifies that the XSLT processor is to apply the `xsl:template`s defined in this XSL document to the current node's children.

Fig. 20.11 | XSL document that transforms `sorting.xml` into XHTML. (Part 1 of 5.)

Outline

```
20    <!-- match book -->
21    <xsl:template match = "book">
22       <head>
23          <title>ISBN <xsl:value-of select = "@isbn"/> -
24             <xsl:value-of select = "title"/></title>
25       </head>
26
27       <body>
28          <h1 style = "color: blue"><xsl:value-of select = "title"/></h1>
29          <h2 style = "color: blue">by
30             <xsl:value-of select = "author/firstName"/>
31             <xsl:text> </xsl:text>
32             <xsl:value-of select = "author/lastName"/>
33          </h2>
34
35          <table style = "border-style: groove; background-color: wheat">
36
37             <xsl:for-each select = "chapters/frontMatter/*">
38                <tr>
39                   <td style = "text-align: right">
40                      <xsl:value-of select = "name()"/>
41                   </td>
```

sorting.xsl

(2 of 5)

Use the `book`'s ISBN (from attribute `isbn`) and the contents of element `title` to create the string that appears in the browser window's title bar.

The `xsl:text` element is used to insert literal text.

Select each element (indicated by an asterisk) that is a child of element `frontMatter`.

Node-set function `name` retrieves the current node's element name.

Fig. 20.11 | XSL document that transforms `sorting.xml` into XHTML. (Part 2 of 5.)

```
42
43                <td>
44                   ( <xsl:value-of select = "@pages"/> pages )
45                </td>
46             </tr>
47          </xsl:for-each>
48
49          <xsl:for-each select = "chapters/chapter">
50             <xsl:sort select = "@number" data-type = "number"
51                order = "ascending"/>
52             <tr>
53                <td style = "text-align: right">
54                   Chapter <xsl:value-of select = "@number"/>
55                </td>
56
57                <td>
58                   <xsl:value-of select = "text()"/>
59                   ( <xsl:value-of select = "@pages"/> pages )
60                </td>
61             </tr>
62          </xsl:for-each>
```

sorting.xsl

(3 of 5)

Use element `xsl:sort` to sort the selected elements by the value given by its `select` attribute.

Use **node-set function** `text` to obtain the text between the `chapter` start and end tags.

Fig. 20.11 | XSL document that transforms `sorting.xml` into XHTML. (Part 3 of 5.)

sorting.xsl

(4 of 5)

```
63
64          <xsl:for-each select = "chapters/appendix">
65             <xsl:sort select = "@number" data-type = "text"
66                order = "ascending"/>
67             <tr>
68                <td style = "text-align: right">
69                   Appendix <xsl:value-of select = "@number"/>
70                </td>
71
72                <td>
73                   <xsl:value-of select = "text()"/>
74                   ( <xsl:value-of select = "@pages"/> pages )
75                </td>
76             </tr>
77          </xsl:for-each>
78       </table>
79
```

Fig. 20.11 | XSL document that transforms `sorting.xml` into XHTML. (Part 4 of 5.)

Outline

```
80          <p style = "color: blue">Pages:
81             <xsl:variable name = "pagecount"
82                select = "sum(chapters//*/@pages)"/>
83             <xsl:value-of select = "$pagecount"/>
84          <br />Media Type: <xsl:value-of select = "media/@type"/></p>
85       </body>
86    </xsl:template>
87 </xsl:stylesheet>
```

sorting.xsl

(5 of 5)

Use an **XSL variable** to store the value of the book's total page count and output the page count to the result tree.

Function `sum` totals a set of values found using XPath.

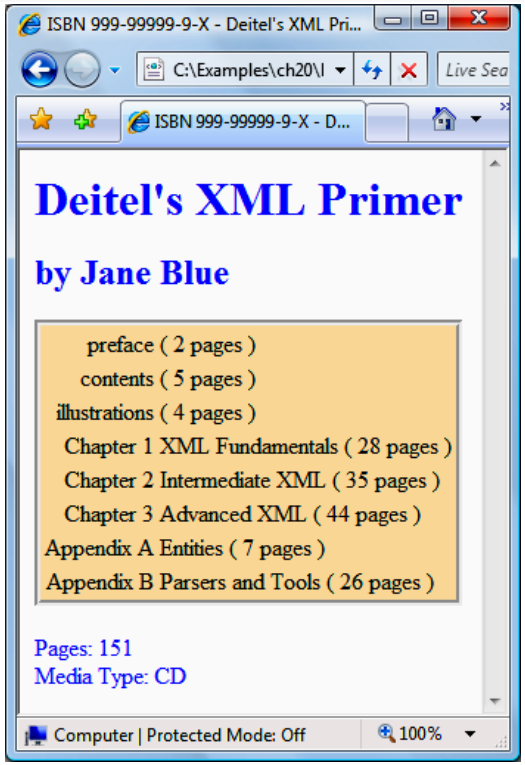

Fig. 20.11 | XSL document that transforms `sorting.xml` into XHTML. (Part 5 of 5.)

20.4 Extensible Stylesheet Language and XSL Transformations (Cont.)

- The `<xsl:apply-templates/>` element specifies that the XSLT processor is to apply the `xsl:templates` defined in this XSL document to the current node's children.
- The `xsl:text` element is used to insert literal text.
- **Node-set function** `name` retrieves the current node's element name.
- Use element `xsl:sort` to sort the selected elements by the value given by its `select` attribute.
- Attribute `data-type`, with value `"number"`, specifies a numeric sort. This attribute also accepts the value `"text"`.

Extensible Stylesheet Language and XSL Transformations (Cont.)

- Attribute `order` can be set to either `"ascending"` or `"descending"`.
- XSL variables cannot be modified after they are initialized.
  - Attribute `name` specifies the variable's name
  - Attribute `select` assigns a value to the variable.
- Function `sum` totals a set of values found using XPath.
- Two slashes in an XPath expression indicate **recursive descent**.

Extensible Stylesheet Language and XSL Transformations (Cont.)

Summary of XSL Style-Sheet Elements

- Figure 20.12 lists commonly used XSL elements.

Element	Description
`<xsl:apply-templates>`	Applies the templates of the XSL document to the children of the current node.
`<xsl:apply-templates match = "`*expression*`">`	Applies the templates of the XSL document to the children of the nodes matching *expression*. The value of the attribute match (i.e., *expression*) must be an XPath expression that specifies elements.
`<xsl:template>`	Contains rules to apply when a specified node is matched.
`<xsl:value-of select =` `"`*expression*`">`	Selects the value of an XML element and adds it to the output tree of the transformation. The required `select` attribute contains an XPath expression.

Fig. 20.12 | XSL style-sheet elements. (Part 1 of 2.)

Extensible Stylesheet Language and XSL Transformations (Cont.)

Element	Description
`<xsl:for-each select = "`*expression*`">`	Applies a template to every node selected by the XPath specified by the `select` attribute.
`<xsl:sort select = "`*expression*`">`	Used as a child element of an `<xsl:apply-templates>` or `<xsl:for-each>` element. Sorts the nodes selected by the `<xsl:apply-template>` or `<xsl:for-each>` element so that the nodes are processed in sorted order.
`<xsl:output>`	Has various attributes to define the format (e.g., XML, XHTML), version (e.g., 1.0, 2.0), document type and MIME type of the output document. MIME types are discussed in Section 22.2. This tag is a top-level element—it can be used only as a child element of an `xsl:stylesheet`.
`<xsl:copy>`	Adds the current node to the output tree.

Fig. 20.12 | XSL style-sheet elements. (Part 2 of 2.)