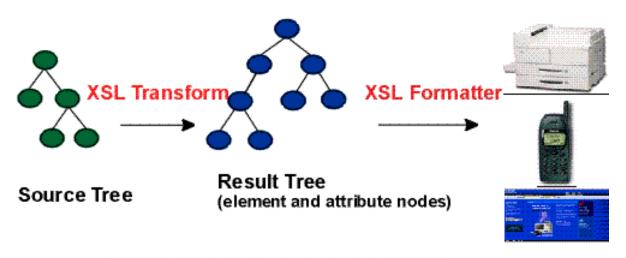


Introduction to XML



XSL (Extensible Stylesheet Language)

- Language for expressing style sheets
- Consists of two parts:
 - Transforming XSLT
 - Formatting XSL-FO



Result XML tree is the result of XSLT processing.



What is XSLT?

- Language for transforming XML documents into other ones
- Designed for use both as part of XSL and independently of XSL
 - E.g., use XSLT to create an output document that contains XSL Formatting Objects, and then feed this output into an XSL-FO processor that generates PDF, for example FOP from Apache (http://xmlgraphics.apache.org/fop/)
- Uses XPath for selecting nodes
 - Locate parts of the source tree document that match templates defined in the XSL stylesheet
- W3C Recommendation

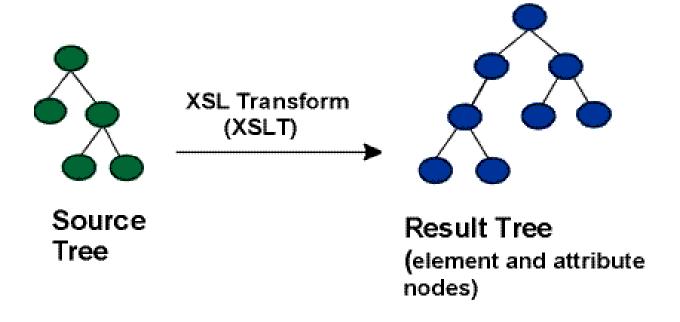


How Tree Transformation Works?

- Style sheet describes rules for transforming a source tree into a result tree
- XSLT processor
 - Reads XML document & XSLT style sheet
 - Outputs a new XML document or fragment
- Transformation is achieved by associating patterns with templates
 - Pattern is matched against elements in source tree
 - Template is instantiated to create part of result tree



With tree transformation, the structure of the result tree can be quite different from the structure of the source tree



In constructing the result tree, the source tree can be filtered and reordered, and arbitrary structure and generated content can be added.

Data Model

- Data model of XSLT is the same as that used by XPath
- Tree structure of XML documents consists of 7 nodes:
 - The root node
 - Element nodes
 - Text nodes
 - Attribute nodes
 - Namespace nodes
 - Processing Instruction nodes
 - Comment nodes

Example: Periodic Table with two Atoms



```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="eg.xsl"?>
<PERIODIC TABLE>
   <ATOM STATE="GAS">
        <NAME>Hydrogen</NAME>
         <SYMBOL>H</SYMBOL>
         <ATOMIC NUMBER>1</ATOMIC NUMBER>
         <ATOMIC WEIGHT>1.00794</ATOMIC WEIGHT>
         <BOILING POINT UNITS="Kelvin">20.28</BOILING POINT>
         <MELTING POINT UNITS="Kelvin">13.81</MELTING POINT>
        <DENSITY UNITS="grams/cubic centimeter">
                 <!-- At 300K, 1 atm -->
                 0.0000899
         </DENSITY>
   </ATOM>
   <ATOM STATE="GAS">
        <NAME>Helium</NAME>
         <SYMBOL>He</SYMBOL>
         <ATOMIC NUMBER>2</ATOMIC NUMBER>
         <ATOMIC WEIGHT>4.0026</ATOMIC WEIGHT>
         <BOILING POINT UNITS="Kelvin">4.216</BOILING POINT>
         <MELTING POINT UNITS="Kelvin">0.95
         <DENSITY UNITS="grams/cubic centimeter">
                 <!-- At 300K -->
                 0.0001785
        </DENSITY>
   </ATOM>
</PERIODIC TABLE>
```



XSL Templates

- Template rules identify the nodes by using a pattern
 - Specified with the xsl:template element
- Content of xsl:template is instantiated
- match attribute is a pattern that identifies the source node or nodes to which the rule applies
 - <xsl:template match = pattern >
 - The syntax for patterns is a subset of XPath expressions



The xsl:apply-templates element



- Tells a processor to compare each child element of the matched source element against the templates in the style sheet
 - If a match is found, output the template for the matched node
 - The template for the matched node may itself contain xsl:apply-templates elements to search for matches for its children

An XSLT style sheet (eg.xsl) for the periodic table



```
<?xml version="1.0"?>
<xsl:stylesheet version="1.0"</pre>
   xmlns:xsl ="http://www.w3.org/1999/XSL/Transform">
        <xsl:template match="/">
                <html>
                   <xsl:apply-templates/> 
                </html>
        </xsl:template>
        <xsl:template match="PERIODIC_TABLE">
                <body>
                   <xsl:apply-templates/>
                </body>
        </xsl:template>
        <xsl:template match="ATOM">
                An Atom
        </xsl:template>
```

</xsl:stylesheet>

The result is:

Computing the value of a node with xsl:value-of





Values of Nodes in xsl:value-of

Node Type:	Value:
Root	The value of the root element
Element	The concatenation of all parsed character data contained in the element, including character data in any of the descendants of the element
Text	The text of the node
Attribute	The attribute value after entities are resolved and leading and trailing white space is stripped; does not include the name of the attribute, the equals sign, or the quotation marks
Namespace	The URI of the namespace
Processing instruction	The data in the processing instruction; does not include the processing instruction, or ?
Comment	The text of the comment, and not included

Processing Multiple Elements with xsl:for-each

 The select="." in this template tells the processor to take the value of the current element, ATOM in this example



Patterns for Matching Nodes

The match attribute of the xsl:template element supports
a complex syntax that allows you to express exactly
which nodes you do and do not want to match

Matching the root node

<xsl:template match="/">



Patterns for Matching Nodes

Matching element names

```
<xsl:template match="ATOM">
```

Matching children with /

```
<xsl:template match="ATOM/SYMBOL"> []
```

<xsl:template match="PERIODIC_TABLE/*/SYMBOL">



Matching descendants with //

```
<xsl:template match="PERIODIC_TABLE//NAME">
<xsl:template match="//ATOMIC_NUMBER">
```

- Matching by ID
 - Used to apply a particular style to a particular single element without changing all other elements of that type
 - id() selector: contains the ID value in single quotes



Matching attributes with @

The output is:

<l>kelvin</l>

<l>kelvin</l>



- Using the or operator
 - Allows a template rule to match multiple patterns

<xsl:template match="ATOMIC_NUMBER|ATOMIC_WEIGHT">

- Testing with []
 - Test for more details about the nodes that match a pattern using []
 - Perform many different tests including:
 - Whether an element contains a given child, attribute, or other node
 - Whether the value of an attribute is a certain string
 - Whether the value of an element matches a string
 - What position a given node occupies in the hierarchy

<xsl:template match="ATOM[MELTING_POINT]">

<xsl:template match="ATOM[DENSITY/@UNITS]">

<xsl:template match="ATOM[ATOMIC_NUMBER='10']">



The Built-in Template Rules

Default rule for elements:

```
<xsl:template match="*|/">
  <xsl:apply-templates/>
</xsl:template>
```

- Applies to element nodes and the root node:
- */ is XPath shorthand for "any element or root node"
- Ensure that all elements are recursively processed



Default rule for text nodes and attributes:

- Matches all text and attribute nodes and outputs the value of the node (<xsl:value-of select="."/>).
- Ensures that at the very least an element's text is output, even if no rule specifically matches it
- Copies attribute values (but not names)



- Default rule for processing instructions and comments:
 - Simply says to do nothing
 - Drop them from the output

<xsl:template match="processing-instruction()|comment()"/>



Attribute Value Templates

- Copy data from input document to attribute values
- Attribute values in curly braces {} takes the place of the

xsl:value-of element

 Ex: To convert the periodic table into empty ATOM elements, the template can be written as:

```
<ATOM NAME="Vanadium"

ATOMIC_WEIGHT="50.9415"

ATOMIC_NUMBER="23" />
```



```
<xsl:template match="ATOM">
          <A HREF="{SYMBOL}.html">
                <xsl:value-of select="NAME"/>
                 </A>
</xsl:template>
```

The output is of the form:

Hydrogen
Helium



Inserting Elements with xsl:element

- Element name is given by an attribute value template in the name attribute of xsl:element
 - Replace ATOM with GAS, LIQUID, and SOLID



Inserting Attributes with xsl:attribute

- Name: name attribute of the xsl:attribute element
- Value: contents of the xsl:attribute element
- Each xsl:attribute element is child of either an xsl:element element or a literal element

```
<xsl:template match="ATOM">
   <LI><A>
      <xsl:attribute name="HREF"> 
          <xsl:value-of select="SYMBOL"/>.html 
      </xsl:attribute>
     <xsl:value-of select="NAME"/>
   </A></LI>
</xsl:template>
The output is of the form:
<LI><A HREF="H.html">Hydrogen</A></LI>
<LI><A HREF="He.html">Helium</A></LI>
```



Generating PI with xsl:processing-instruction

- Target: required name attribute
- Content: contents of the xsl:processing-instruction
 - Replaces PROGRAM elements with a gcc processing instruction:



Generating Comments with xsl:comment

- xsl:comment element
 - no attributes
 - contents are the text of the comment

<xsl:template match="ATOM">

<xsl:comment>There was an atom here once. </xsl:comment>

</xsl:template>

Output: Replaces ATOM nodes with this comment:

<!--There was an atom here once.-->



Generating Text with xsl:text

Inserts its contents into the output document

<xsl:template match="ATOM">
 <xsl:text>There was an atom here once.</xsl:text>
</xsl:template>

Output: replaces each ATOM element with the string:

There was an atom here once.

Copying the Context Node with xsl:copy

Copies source node into output tree

</xsl:stylesheet>

- Child elements, attributes, and other content are not automatically copied
 - Ex: strips comments from a document

```
<?xml version="1.0"?>
<xsl:stylesheet version="1.0"
    xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
    <xsl:template match="*|@*|processing-instruction()|text()">
    <xsl:copy>
    <xsl:apply-templates select="*|@*|processing-instruction()|text()"/>
    </xsl:copy>
    </xsl:template>
```



Sorting Output Elements

- xsl:sort element
 - Sorts output elements into a different order than they appear in the input
 - Appears as a child of xsl:apply-templates or xsl:for-each
 - The select attribute of xsl:sort defines the key used to sort the element's output

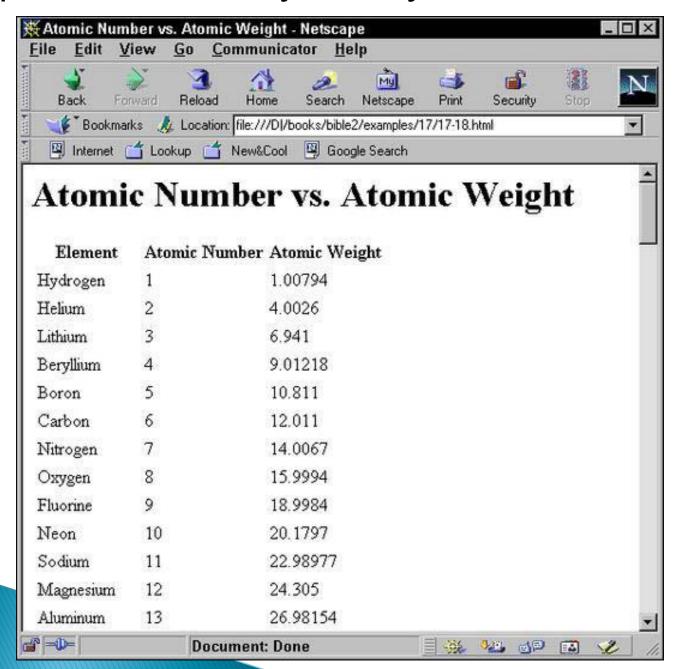
Ex: An XSLT style sheet that sorts by atomic number



```
<?xml version="1.0"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
   <xsl:template match="PERIODIC_TABLE">
      <html>
         <head>
          <title>Atomic Number vs. Atomic Weight</title>
         </head>
         <body>
           <h1>Atomic Number vs. Atomic Weight</h1>
           Element
             Atomic Number
             Atomic Weight
             <xsl:apply-templates>
                   <xsl:sort data-type="number"</pre>
                              select="ATOMIC_NUMBER" order="ascending"/>
             </xsl:apply-templates>
           </body>
     </html>
   </xsl:template>
   <xsl:template match="ATOM">
         <xsl:apply-templates select="NAME"/>
           <xsl:apply-templates select="ATOMIC_NUMBER"/>
           <xsl:apply-templates select="ATOMIC_WEIGHT"/>
         </xsl:template>
</xsl:stylesheet>
```

Output: Atoms numerically sorted by atomic number







Modes

Modes allow an element to be processed multiple times,
 each time producing a different result

Both xsl:template and xsl:apply-templates have an optional mode attribute

Ex: An XSLT style sheet that uses modes to format the same data differently in two different p

```
<?xml version="1.0"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
    <xsl:template match="/PERIODIC_TABLE">
       <HTML>
          <HEAD>
             <TITLE>The Elements</TITLE>
          </HEAD>
          <BODY>
              <H2>Table of Contents</H2>
              <UL>
                <xsl:apply-templates select="ATOM" mode="toc"/>
              </UL>
              <H2>The Elements</H2>
                     </BODY> <xsl:apply-templates select="ATOM" mode="full"/>
       </HTML>
    </xsl:template>
    <xsl:template match="ATOM" mode="toc">
          <LI><A>
             <xsl:attribute name="HREF">#<xsl:value-of select="SYMBOL"/></xsl:attribute>
              <xsl:value-of select="NAME"/>
          </A></LI>
   </xsl:template>
    <xsl:template match="ATOM" mode="full">
          <H3><A>
              <xsl:attribute name="NAME"> <xsl:value-of select="SYMBOL"/> </xsl:attribute>
              <xsl:value-of select="NAME"/>
          </A></H3>
          <P> <xsl:value-of select="."/> </P>
   </xsl:template>
</xsl:stylesheet>
```



Defining Constants with xsl:variable

- xsl:variable element defines a named string for use elsewhere in the style sheet via an attribute value template
 - Single attribute, name, which provides the variable name
 - Contents provide the replacement text

```
<xsl:variable name="copy01">
    Copyright 2001 Elliotte Rusty Harold
</xsl:variable>
```

```
Accessing the variable "copy01" 

<BLOCK COPYRIGHT="{$copy01}"> </BLOCK> □
```



Named Templates

- xsl:template element has a name attribute by which it can be explicitly invoked
- xsl:call-template element has a required name argument that names the called template
- When processed, xsl:call-template is replaced by the contents of the xsl:template element it names



```
<xsl:template name="ATOM_CELL">
```

</xsl:template>

Ex: Calling the named template **ATOM_CELL**



Passing Parameters to Templates

- xsl:with-param element
 - The value of the parameter is specified in the same way as for xsl:variable and xsl:param

Ex: Passing the parameter "file"



```
The Output is of the form:

<font face="Times, serif" color="blue" size="2">
<b>
<b>
<a href="atomic_number.html">52</a>
</b>
</font>
```



Making Choices

- xsl:if
 - Provides a simple facility for changing output based on a pattern
 - test attribute of xsl:if contains an expression that evaluates to a boolean

Ex: A comma and a space is added after all names except the last name



xsl:choose

- Selects one of several possible outputs
 - xsl:when: each condition and its associated output template
 - If none of the xsl:when elements are true, the xsl:otherwise child element is instantiated



Ex: Changes the color of the output based on whether the STATE attribute of the ATOM element is SOLID or LIQUID

```
<xsl:template match="ATOM">
  <xsl:choose>
   <xsl:when test="@STATE='SOLID'"> 
        <P style="color: black"> <xsl:value-of select="."/> </P>
   </xsl:when>
   <xsl:when test="@STATE='LIQUID'">
        <P style="color: blue"> <xsl:value-of select="."/> </P>
   </xsl:when>
   <xsl:otherwise>
        <P style="color: green"> <xsl:value-of select="."/> </P>
   </xsl:otherwise>
  </xsl:choose>
</xsl:template>
```



Output Methods

- xsl:output element
 - Allows stylesheet authors to specify how they wish the result tree to be output
 - method attribute specifies which output method to use and normally has one of the three values (XML, HTML, Text)

```
<!-- Category: top-level-element -->
<xsl:output method = "xml" | "html" | "text" />
```



- XSLT does not analyze every node of the source tree
 - Selectively navigates the source tree using XPath's select and match attributes
- For XSLT to function, the source tree must be properly structured
- XSL style sheets can be connected directly to an XML document by adding an xml-stylesheet processing instruction to the XML document



- Two tree structures are involved in transforming an XML document using XSLT
 - source tree (the document being transformed)
 - result tree (the result of the transformation)
- XPath character / (a forward slash)
 - Selects the document root
 - In XPath, a leading forward slash specifies that we are using absolute addressing
 - An XPath expression with no beginning forward slash uses relative addressing



```
<?xml version = "1.0"?>
    <?xml-stylesheet type = "text/xsl" href = "sports.xsl"?>
 2
 3
    <!-- Fig. 15.18: sports.xml -->
 5
    <!-- Sports Database -->
 6
 7
    <sports>
       <game id = "783">
8
9
           <name>Cricket</name>
10
           <paragraph>
11
              More popular among commonwealth nations.
12
13
           </paragraph>
       </game>
14
15
16
       <qame id = "239">
           <name>Baseball</name>
17
18
           <paragraph>
19
20
              More popular in America.
           </paragraph>
21
22
       </game>
23
```

Fig. 15.18 | XML document that describes various sports. (Part 1 of 2.)



Fig. 15.18 | XML document that describes various sports. (Part 2 of 2.)



```
<?xml version = "1.0"?> 
     <!-- Fig. 15.19: sports.xsl -->
 2
     <!-- A simple XSLT transformation -->
 3
 4
 5
     <!--preference XSL style sheet URI -->
     <xsl-stylesheet version = "1.0"</pre>
        xmlns:xsl = "http://www.w3.org/1999/XSL/Transform">
 7
 8
 9
        <xsl:output method = "html" doctype-system = "about:legacy-compat" />
        <xsl:template match = "/"> <!-- match root element -->
10
11
        <html xmlns = "http://www.w3.org/1999/xhtml">
12
           <head>
13
              <meta charset = "utf-8"/>
14
              <link rel = "stylesheet" type = "text/css" href = "style.css"/>
15
16
              <title>Sports</title>
           </head>
17
18
19
           <body>
              20
              <caption>Information about various sports</caption>
21
                 <thead>
22
23
```

Fig. 15.19 | XSLT that creates elements and attributes in an HTML5 document. (Part 1 of 2.)



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

Fig. 15.19 | XSLT that creates elements and attributes in an HTML5 document. (Part 2 of 2.)



```
<?xml version = "1.0"?>
 2
    <?xml-stylesheet type = "text/xsl" href = "sorting.xsl"?>
 3
    <!-- Fig. 15.20: sorting.xml -->
4
    <!-- XML document containing book information -->
 5
    <book isbn = "999-99999-9-X">
       <title>Deitel&apos;s XML Primer</title>
 7
8
       <author>
10
          <firstName>Jane</firstName>
11
          <lastName>Blue</lastName>
12
       </author>
13
       <chapters>
14
          <frontMatter>
15
             16
             <contents pages = "5" />
17
             <illustrations pages = "4" />
18
          </frontMatter>
19
20
```

Fig. 15.20 | XML document containing book information. (Part 1 of 2.)



```
<chapter number = "3" pages = "44">Advanced XML</chapter>
21
          <chapter number = "2" pages = "35">Intermediate XML</chapter>
22
23
          <appendix number = "B" pages = "26">Parsers and Tools</appendix>
          <appendix number = "A" pages = "7">Entities</appendix>
24
          <chapter number = "1" pages = "28">XML Fundamentals
25
       </chapters>
26
27
       <media type = "CD" />
28
    </book>
29
```

Fig. 15.20 | XML document containing book information. (Part 2 of 2.)



```
<?xml version = "1.0"?> \overline{\circ}
2
3
    <!-- Fig. 15.21: sorting.xsl -->
    <!-- Transformation of book information into HTML5 -->
    <xsl:stylesheet version = "1.0"</pre>
       xmlns:xsl = "http://www.w3.org/1999/XSL/Transform">
7
8
       <!-- write XML declaration and DOCTYPE DTD information -->
9
       <xsl:output method = "html" doctype-system = "about:legacy-compat" />
10
11
       <!-- match document root -->
       <xsl:template match = "/">
12
           <html>
13
              <xsl:apply-templates/>
14
           </html>
15
       </xsl:template>
16
17
```

Fig. 15.21 | XSL document that transforms sorting.xml into HTML5. (Part I of 5.)



```
18
       <!-- match book -->
       <xsl:template match = "book">
19
          <head>
20
21
             <meta charset = "utf-8"/>
             <link rel = "stylesheet" type = "text/css" href = "style.css"/>
22
             <title>ISBN <xsl:value-of select = "@isbn"/> -
23
                <xsl:value-of select = "title"/></title>
24
25
          </head>
26
27
          <body>
             <h1><xs1:value-of select = "title"/></h1>
28
29
             <h2>by
30
                <xsl:value-of select = "author/lastName"/>,
                <xsl:value-of select = "author/firstName"/></h2>
31
32
33
             34
35
                <xsl:for-each select = "chapters/frontMatter/*">
36
                   37
                         <xsl:value-of select = "name()"/>
38
                      39
40
```

Fig. 15.21 | XSL document that transforms sorting.xml into HTML5. (Part 2 of 5.)



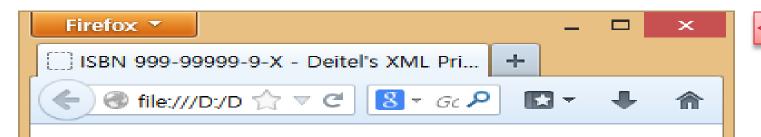
```
41
                        ( <xsl:value-of select = "@pages"/> pages )
42
43
                     44
               </xsl:for-each>
45
46
               <xsl:for-each select = "chapters/chapter">
47
                  <xsl:sort select = "@number" data-type = "number"</pre>
48
                       order = "ascending"/>
49
50
                  >
51
                     52
                        Chapter <xsl:value-of select = "@number"/>
53
                     54
55
                     56
                        <xsl:value-of select = "text()"/>
                        ( <xsl:value-of select = "@pages"/> pages )
57
58
                     59
               </xsl:for-each>
60
61
```

Fig. 15.21 | XSL document that transforms sorting.xml into HTML5. (Part 3 of 5.)



```
62
                <xsl:for-each select = "chapters/appendix">
                   <xsl:sort select = "@number" data-type = "text"</pre>
63
                      order = "ascending"/>
64
65
                   66
                      Appendix <xsl:value-of select = "@number"/>
67
68
                      69
                      70
                         <xsl:value-of select = "text()"/>
71
                         ( <xsl:value-of select = "@pages"/> pages )
72
73
                      74
                   </xsl:for-each>
75
76
             77
             Pages:
78
                <xsl:variable name = "pagecount"</pre>
79
                   select = "sum(chapters//*/@pages)"/>
80
                <xsl:value-of select = "$pagecount"/> \rightarrow
81
             Media Type: <xsl:value-of select = "media/@type"/>
82
83
          </body>
       </xsl:template>
84
85
    </xsl:stylesheet>
```

Fig. 15.21 | XSL document that transforms sorting.xml into HTML5. (Part 4 of 5.)



Deitel's XML Primer

by Blue, Jane

```
preface ( 2 pages )
contents ( 5 pages )
illustrations ( 4 pages )
Chapter 1 XML Fundamentals ( 28 pages )
Chapter 2 Intermediate XML ( 35 pages )
Chapter 3 Advanced XML ( 44 pages )
Appendix A Entities ( 7 pages )
Appendix B Parsers and Tools ( 26 pages )
```

Pages: 151

Media Type: CD