

# Contents

<b>Introduction</b>	<b>xxv</b>
<b>Part I: Introduction</b>	<b>1</b>
<b>Chapter 1: What Is XML?</b>	<b>3</b>
<b>Of Data, Files, and Text</b>	<b>3</b>
Binary Files	4
Text Files	5
A Brief History of Markup	6
<b>So What Is XML?</b>	<b>7</b>
What Does XML Buy Us?	10
HTML and XML: Apples and Red Delicious Apples	13
Hierarchies of Information	14
What's a Document Type?	17
<b>What Is the World Wide Web Consortium?</b>	<b>18</b>
What Are the Pieces That Make Up XML?	19
<b>Where Is XML Used, and Where Can It Be Used?</b>	<b>20</b>
Reducing Server Load	20
Web Site Content	20
Remote Procedure Calls	20
e-Commerce	20
<b>Summary</b>	<b>21</b>
<b>Exercise Questions</b>	<b>21</b>
Question 1	21
Question 2	21
<b>Chapter 2: Well-Formed XML</b>	<b>23</b>
<b>Parsing XML</b>	<b>23</b>
<b>Tags and Text and Elements, Oh My!</b>	<b>24</b>
Rules for Elements	30
<b>Attributes</b>	<b>38</b>
Why Use Attributes?	41
<b>Comments</b>	<b>44</b>
<b>Empty Elements</b>	<b>47</b>

## Contents

---

<b>XML Declaration</b>	<b>49</b>
Encoding	49
Standalone	52
<b>Processing Instructions</b>	<b>54</b>
Is the XML Declaration a Processing Instruction?	55
<b>Illegal PCDATA Characters</b>	<b>57</b>
Escaping Characters	58
CDATA Sections	59
<b>Errors in XML</b>	<b>62</b>
<b>Summary</b>	<b>63</b>
<b>Exercise Questions</b>	<b>63</b>
Question 1	64
Question 2	64
<b>Chapter 3: XML Namespaces</b>	<b>65</b>
<b>Why Do We Need Namespaces?</b>	<b>65</b>
Using Prefixes	66
So Why Doesn't XML Just Use These Prefixes?	68
<b>How XML Namespaces Work</b>	<b>69</b>
Default Namespaces	72
Do Different Notations Make Any Difference?	78
Namespaces and Attributes	80
<b>What Exactly Are URIs?</b>	<b>82</b>
URLs	83
URNs	84
Why Use URLs for Namespaces, Not URNs?	84
What Do Namespace URIs Really Mean?	85
RDDL	85
<b>When Should I Use Namespaces?</b>	<b>86</b>
<b>Summary</b>	<b>87</b>
<b>Exercise Questions</b>	<b>87</b>
Question 1	87
Question 2	87
Question 3	88
<b>Part II: Validation</b>	<b>89</b>
<b>Chapter 4: Document Type Definitions</b>	<b>91</b>
<b>Running the Samples</b>	<b>92</b>

## Contents

---

Preparing the Ground	92
The Document Type Declaration	97
<b>Sharing Vocabularies</b>	<b>101</b>
<b>Anatomy of a DTD</b>	<b>102</b>
Element Declarations	103
Attribute Declarations	117
Entities	129
Notation Declarations	142
<b>Developing DTDs</b>	<b>145</b>
<b>DTD Limitations</b>	<b>146</b>
DTD Syntax	146
XML Namespaces	146
Data Typing	146
Limited Content Model Descriptions	147
<b>Summary</b>	<b>147</b>
<b>Exercise Questions</b>	<b>147</b>
Question 1	147
Question 2	148
Question 3	148
<b>Chapter 5: XML Schemas</b>	<b>149</b>
<b>Benefits of XML Schemas</b>	<b>150</b>
XML Schemas Use XML Syntax	150
XML Schema Namespace Support	150
XML Schema Data Types	150
XML Schema Content Models	151
<b>Do We Still Need DTDs?</b>	<b>151</b>
<b>XML Schemas</b>	<b>151</b>
The XML Schema Document	152
Running the Samples	152
<schema> Declarations	157
<element> Declarations	159
<complexType> Declarations	169
<group> Declarations	170
Content Models	172
<attribute> Declarations	182
<attributeGroup> Declarations	188
Data Types	192
<simpleType> Declarations	200
Creating Elements with <simpleType> Content and Attributes	206

## Contents

---

<b>Creating a Schema from Multiple Documents</b>	<b>207</b>
<import> Declarations	208
<include> Declarations	212
<b>Documenting XML Schemas</b>	<b>216</b>
Comments	216
Attributes from Other Namespaces	217
Annotations	218
<b>Summary</b>	<b>219</b>
<b>Exercise Questions</b>	<b>220</b>
Question 1	220
Question 2	220
Question 3	220
<b>Chapter 6: RELAX NG</b>	<b>221</b>
<b>Basic RELAX NG Patterns</b>	<b>222</b>
Introduction to Patterns and Element Patterns	222
Attribute Patterns	225
Enumerated Values	227
Co-Occurrence Constraints	228
Mixed Content Pattern	229
The Empty Pattern	229
Connector Patterns and Grouping	230
The Interleave Pattern	231
<b>Combining and Reusing Patterns and Grammars</b>	<b>234</b>
Named Patterns	234
assignMethod and Combining Named Pattern Definitions	237
Schema Modularization Using the Include Directive	238
Redefining Included Named Patterns	239
Removing Patterns with the notAllowed Pattern	240
Extending Schemas with Composition	241
Extensibility in RELAX NG	242
Nested Grammars and the RELAX NG Grammar Pattern	243
<b>Additional RELAX NG Features</b>	<b>247</b>
Namespaces	247
Name-Classes	249
Data Types	252
List Patterns	254
Comments and Divisions	255
<b>Useful Resources</b>	<b>255</b>
<b>Summary</b>	<b>256</b>

## Contents

---

<b>Exercise Questions</b>	<b>256</b>
Question 1	256
Question 2	256
<b>Part III: Processing</b>	<b>257</b>
<b>Chapter 7: XPath</b>	<b>259</b>
<b>Ways of Looking at an XML Document</b>	<b>260</b>
<b>Serialized Document</b>	<b>260</b>
XPath Data Model	261
The Document Object Model	261
The XML Information Set	261
<b>Visualizing XPath</b>	<b>261</b>
Context	262
What Is a Node?	263
XPath 1.0 Types	266
<b>Abbreviated and Unabbreviated Syntax</b>	<b>267</b>
<b>XPath 1.0 Axes</b>	<b>268</b>
Child Axis	269
attribute Axis	270
ancestor Axis	272
ancestor-or-self Axis	273
descendant Axis	273
descendant-or-self Axis	274
following Axis	274
following-sibling Axis	276
namespace Axis	276
parent Axis	279
preceding Axis	279
preceding-sibling Axis	280
self Axis	281
<b>XPath 1.0 Functions</b>	<b>282</b>
Boolean Functions	282
Node-Set Functions	283
Numeric Functions	283
String Functions	284
<b>Predicates</b>	<b>284</b>
<b>Structure of XPath Expressions</b>	<b>285</b>
<b>Looking Forward to XPath 2.0</b>	<b>287</b>
Revised XPath Data Model	288

## Contents

---

W3C XML Schema Data Types	288
Additional XPath 2.0 Functions	288
<b>Summary</b>	<b>288</b>
<b>Exercise Questions</b>	<b>288</b>
Question 1	289
Question 2	289
<b>Chapter 8: XSLT</b>	<b>291</b>
<b>What Is XSLT?</b>	<b>291</b>
Restructuring XML	292
Presenting XML Content	292
<b>How Does an XSLT Processor Work?</b>	<b>292</b>
<b>Running the Examples</b>	<b>293</b>
Introducing the Saxon XSLT Processor	293
Installing the Saxon XSLT Processor	293
<b>Procedural versus Declarative Programming</b>	<b>295</b>
Procedural Programming	296
Declarative Programming	296
<b>Foundational XSLT Elements</b>	<b>297</b>
The <xsl:stylesheet> Element	299
The <xsl:template> Element	300
The <xsl:apply-templates> Element	300
<b>Getting Information from the Source Tree</b>	<b>301</b>
The <xsl:value-of> Element	301
The <xsl:copy> Element	303
The <xsl:copy-of> Element	307
<b>Influencing the Output with the &lt;xsl:output&gt; Element</b>	<b>309</b>
<b>Conditional Processing</b>	<b>310</b>
The <xsl:if> Element	310
The <xsl:choose> Element	311
<b>The &lt;xsl:for-each&gt; Element</b>	<b>314</b>
<b>The &lt;xsl:sort&gt; Element</b>	<b>316</b>
<b>XSLT Modes</b>	<b>318</b>
<b>XSLT Variables and Parameters</b>	<b>323</b>
<b>Named Templates and the &lt;xsl:call-template&gt; Element</b>	<b>325</b>
<b>XSLT Functions</b>	<b>325</b>
<b>Looking Forward to XSLT 2.0</b>	<b>326</b>
<b>Summary</b>	<b>326</b>
<b>Exercise Questions</b>	<b>326</b>
Question 1	327
Question 2	327

---

<b>Part IV: Databases</b>	<b>329</b>
<b>Chapter 9: XQuery, the XML Query Language</b>	<b>331</b>
<b>Why XQuery?</b>	<b>332</b>
Historical Factors	332
Technical Factors	333
Current Status	333
<b>XQuery Tools</b>	<b>335</b>
Saxon	336
X-Hive.com Online	337
X-Hive Database	338
Tamino Database	338
SQL Server 2005	338
Oracle	338
<b>Some XQuery Examples</b>	<b>338</b>
Input Functions	339
Retrieving Nodes	340
Element Constructors	343
The XQuery Prolog	347
Computed Constructors	349
Syntax	351
<b>The XQuery Data Model</b>	<b>352</b>
Shared Data Model with XPath 2.0 and XSLT 2.0	352
Node Kinds	352
Sequences of Node-Sets	352
Document Order	353
Comparing Items and Nodes	353
Types in XQuery	353
Axes in XQuery	353
<b>XQuery Expressions</b>	<b>353</b>
FLWOR Expressions	353
<b>XQuery Functions</b>	<b>360</b>
The concat() Function	360
The count() Function	361
<b>Using Parameters with XQuery</b>	<b>361</b>
<b>Proprietary Extensions to XQuery</b>	<b>362</b>
Insert Functionality	362
Delete Functionality	363
Update Functionality	363

## Contents

---

<b>Looking Ahead</b>	<b>363</b>
Update Functionality (as Part of XQuery)	363
Full-Text Search	364
<b>Summary</b>	<b>364</b>
<b>Exercise Questions</b>	<b>364</b>
Question 1	364
Question 2	364
<b>Chapter 10: XML and Databases</b>	<b>365</b>
<b>The Need for Efficient XML Data Stores</b>	<b>365</b>
Increase in Amount of XML	366
Comparing XML-Based Data and Relational Data	367
<b>Approaches to Storing XML</b>	<b>368</b>
Producing XML from Relational Databases	368
Moving XML to Relational Databases	368
Native XML Databases	368
<b>Using Native XML Databases</b>	<b>369</b>
Obtaining and Installing Xindice	370
Adding Information	374
Retrieving Information	378
Searching for Information	379
Programming Xindice	380
<b>XML in RDBMS</b>	<b>380</b>
XML Functionality in SQL Server 2000	381
XML Functionality in SQL Server “Yukon”	389
<b>Choosing a Database to Store XML</b>	<b>399</b>
<b>Looking Ahead</b>	<b>399</b>
<b>Summary</b>	<b>399</b>
<b>Exercise Questions</b>	<b>400</b>
Question 1	400
Question 2	400
<b>Part V: Programming</b>	<b>401</b>
<b>Chapter 11: The XML Document Object Model (DOM)</b>	<b>403</b>
<b>What Is an XML Document Object Model For?</b>	<b>403</b>
Interfaces and Objects	404

## Contents

---

<b>The Document Object Model at the W3C</b>	<b>405</b>
XML DOM Implementations	406
<b>Two Ways to View DOM Nodes</b>	<b>406</b>
<b>Overview of the XML DOM</b>	<b>407</b>
<b>Tools to Run the Examples</b>	<b>408</b>
<b>The Node Object</b>	<b>412</b>
Properties of the Node Object	413
Methods of the Node Object	417
Loading an XML Document	418
The NamedNodeMap Object	427
The NodeList Object	432
The DOMException Object	432
<b>The Document Interface</b>	<b>434</b>
<b>How the XML DOM Is Used in InfoPath 2003</b>	<b>436</b>
<b>Summary</b>	<b>437</b>
<b>Exercise Questions</b>	<b>438</b>
Question 1	438
Question 2	438
<b>Chapter 12: Simple API for XML (SAX)</b>	<b>439</b>
<b>What Is SAX and Why Was It Invented?</b>	<b>439</b>
A Simple Analogy	440
A Brief History of SAX	440
<b>Where to Get SAX</b>	<b>441</b>
Setting Up SAX	442
<b>Receiving SAX Events</b>	<b>442</b>
ContentHandler Interface	443
ErrorHandler Interface	462
DTDHandler Interface	467
EntityResolver Interface	467
Features and Properties	468
Extension Interfaces	472
<b>Good SAX and Bad SAX</b>	<b>474</b>
<b>Consumers, Producers, and Filters</b>	<b>475</b>
<b>Other Languages</b>	<b>475</b>
<b>Summary</b>	<b>476</b>
<b>Exercise Questions</b>	<b>476</b>
Question 1	477
Question 2	477

## Contents

---

<b>Part VI: Communication</b>	<b>479</b>
<b>Chapter 13: RSS and Content Syndication</b>	<b>481</b>
<b>Syndication and Meta Data</b>	<b>481</b>
Syndication Systems	482
The Origin of RSS Species	484
RSS-DEV and RSS 1.0	489
Userland and RSS 2.0	492
Atom	494
<b>Working with News Feeds</b>	<b>496</b>
Newsreaders	497
Data Quality	497
<b>A Simple Aggregator</b>	<b>498</b>
Modeling Feeds	498
Program Flow	501
Implementation	501
Transforming RSS with XSLT	518
<b>Useful Resources</b>	<b>528</b>
<b>Summary</b>	<b>529</b>
<b>Exercise Questions</b>	<b>530</b>
Question 1	530
Question 2	530
<b>Chapter 14: Web Services</b>	<b>531</b>
<b>What Is an RPC?</b>	<b>531</b>
<b>RPC Protocols</b>	<b>533</b>
DCOM	533
IIOP	534
Java RMI	535
<b>The New RPC Protocol: Web Services</b>	<b>535</b>
XML-RPC	536
The Network Transport	539
<b>Taking a REST</b>	<b>549</b>
<b>The Web Services Stack</b>	<b>552</b>
SOAP	552
WSDL	553
UDDI	554
Surrounding Specifications	555
<b>Summary</b>	<b>557</b>

## Contents

---

<b>Exercise Questions</b>	<b>557</b>
Question 1	557
Question 2	558
<b>Chapter 15: SOAP and WSDL</b>	<b>559</b>
<b>Laying the Groundwork</b>	<b>559</b>
<b>The New RPC Protocol: SOAP</b>	<b>562</b>
Just RESTing	568
<b>Basic SOAP Messages</b>	<b>568</b>
<Envelope>	569
<Body>	570
Encoding Style	570
<b>More Complex SOAP Interactions</b>	<b>576</b>
<Header>	576
<Fault>	578
<b>Defining Web Services: WSDL</b>	<b>587</b>
<definitions>	588
Other Bindings	596
<b>Summary</b>	<b>598</b>
<b>Exercise Questions</b>	<b>599</b>
Question 1	599
Question 2	599
<b>Part VII: Display</b>	<b>601</b>
<b>Chapter 16: XHTML</b>	<b>603</b>
<b>Separating Style from Content</b>	<b>604</b>
<b>Learning XHTML 1.0</b>	<b>605</b>
The Strict, Transitional, and Frameset DTDs of XHTML 1.0	606
Basic Changes in Writing XHTML	607
Stricter Documents Make Faster and Lighter Processors	618
XHTML Tools	618
Validating XHTML Documents	619
Validation Pitfalls	620
<b>Modularized XHTML</b>	<b>622</b>
Module Implementations	624
XHTML 1.1	624
XHTML Basic	625
<b>Summary</b>	<b>626</b>

## Contents

---

<b>Exercise Questions</b>	<b>627</b>
Question 1	628
Question 2	628
<b>Chapter 17: Cascading Style Sheets (CSS)</b>	<b>629</b>
<b>Why Style Sheets?</b>	<b>630</b>
<b>Introducing CSS</b>	<b>631</b>
CSS Properties	632
Inheritance	633
<b>Using CSS with XHTML versus Other XML Vocabularies</b>	<b>637</b>
<b>Attaching the Style Sheet to an XML Document</b>	<b>637</b>
<b>Selectors</b>	<b>638</b>
<b>Using CSS for Layout of XML Documents</b>	<b>639</b>
Understanding the Box Model	640
Positioning in CSS	644
<b>Laying Out Tabular Data</b>	<b>656</b>
<b>Links in XML Documents</b>	<b>658</b>
XLink Support in Netscape	659
Forcing Links Using the HTML Namespace	662
<b>Images in XML Documents</b>	<b>662</b>
<b>Using CSS to Add Content to Documents</b>	<b>663</b>
<b>Attribute Content</b>	<b>666</b>
Attribute Selectors	666
Using Attribute Values in Documents	666
<b>Summary</b>	<b>668</b>
<b>Exercise Questions</b>	<b>669</b>
Question 1	669
Question 2	670
Question 3	670
Question 4	670
<b>Chapter 18: Scalable Vector Graphics (SVG)</b>	<b>671</b>
<b>What Is SVG?</b>	<b>671</b>
Scalable, Vector, Graphics	672
Putting SVG to Work	672
An SVG Toolkit	673
<b>Getting Started</b>	<b>674</b>
The Painter's Model	677
Grouping	679

## Contents

---

Transformations	679
Paths	680
Images	683
Text	683
Comments, Annotation, and Meta Data	684
Scripting	686
SVG on Your Website	688
<b>Tangram: A Simple Application</b>	<b>688</b>
XHTML Wrapper	689
SVG Shapes	690
Tangram Script	694
<b>The SVG Specification</b>	<b>702</b>
<b>Useful Resources</b>	<b>706</b>
<b>Summary</b>	<b>706</b>
<b>Exercise Questions</b>	<b>707</b>
Question 1	707
Question 2	708
<b>Chapter 19: XForms</b>	<b>709</b>
<b>How Does XForms Improve on HTML Forms?</b>	<b>709</b>
<b>XForms Tools</b>	<b>710</b>
XForms Viewers	710
XForms Designers	713
<b>An Illustrative XForms Example</b>	<b>714</b>
The XForms Model	716
XML Namespaces in XForms Documents	721
<b>XForms Form Controls</b>	<b>722</b>
The xforms:input Element	722
The xforms:secret Element	723
The xforms:textarea Element	723
The xforms:output Element	723
The xforms:upload Element	724
The xforms:range Element	724
The xforms:trigger Element	724
The xforms:submit Element	724
The xforms:select Element	725
The xforms:select1 Element	726
<b>XForms Model Item Properties</b>	<b>732</b>
The xforms:bind Element	732
<b>XForms Events</b>	<b>737</b>

## Contents

---

<b>The XForms Action Module</b>	<b>737</b>
The xforms:action Element	738
The xforms:dispatch Element	738
The xforms:load Element	738
The xforms:message Element	738
The xforms:rebuild Element	739
The xforms:recalculate Element	739
The xforms:refresh Element	739
The xforms:reset Element	739
The xforms:revalidate Element	739
The xforms:send Element	740
The xforms:setfocus Element	740
The xforms:setvalue Element	740
Developing and Debugging XForms	740
<b>Commercial Alternatives to XForms</b>	<b>741</b>
Microsoft InfoPath 2003	741
Adobe XML/PDF Forms Designer	742
<b>Summary</b>	<b>743</b>
<b>Exercise Questions</b>	<b>743</b>
Question 1	743
Question 2	743
<b>Part VIII: Case Studies</b>	<b>745</b>
<b>Chapter 20: Case Study 1: .NET XML Web Services</b>	<b>747</b>
<b>The XML Web Service</b>	<b>747</b>
<b>Consuming the Web Service</b>	<b>757</b>
The Web Page Client	757
The .NET Client	765
<b>Summary</b>	<b>768</b>
<b>Chapter 21: Case Study 2: XML and PHP</b>	<b>771</b>
<b>Elementary Programming in PHP</b>	<b>772</b>
<b>Basic Information about PHP</b>	<b>773</b>
Programming Principles	773
Constants and Variables	777
Operators	779
Getting Data from the Browser Screen	782
Getting Data from a File	785

---

Contents

<b>Building a Program</b>	<b>787</b>
<b>Useful Resources</b>	<b>797</b>
<b>Summary</b>	<b>798</b>
<b>Part IX: Appendixes</b>	<b>799</b>
<b>Appendix A: Exercise Solutions</b>	<b>801</b>
<b>Appendix B: The XML Document Object Model</b>	<b>857</b>
<b>Appendix C: XPath 1.0 Reference</b>	<b>871</b>
<b>Appendix D: XSLT 1.0 Reference</b>	<b>883</b>
<b>Appendix E: XML Schema Element and Attribute Reference</b>	<b>911</b>
<b>Appendix F: Schema Data Types Reference</b>	<b>941</b>
<b>Index</b>	<b>961</b>

