

# Contents

<b>Acknowledgments</b>	<b>ix</b>
<b>Introduction</b>	<b>xxvii</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
No, Really — What's a Document Type?	18
<b>Origin of the XML Standards</b>	<b>18</b>
What Is the World Wide Web Consortium?	18
Components of XML	19
<b>Where XML Can Be Used, and What You Can Use It For</b>	<b>20</b>
Reducing Server Load	20
Website Content	20
Distributed Computing	21
e-Commerce	21
<b>Summary</b>	<b>22</b>
<b>Exercise Questions</b>	<b>22</b>
Question 1	22
Question 2	22
<b>Chapter 2: Well-Formed XML</b>	<b>23</b>
<b>Parsing XML</b>	<b>24</b>
<b>Tags and Text and Elements, Oh My!</b>	<b>24</b>
Rules for Elements	31

# Contents

---

<b>Attributes</b>	<b>39</b>
When to Use Attributes	43
<b>Comments</b>	<b>45</b>
<b>Empty Elements</b>	<b>49</b>
<b>XML Declarations</b>	<b>50</b>
Version	51
Encoding	51
Standalone	53
<b>Processing Instructions</b>	<b>56</b>
<b>Illegal PCDATA Characters</b>	<b>59</b>
Escaping Characters	60
CDATA Sections	61
<b>Errors in XML</b>	<b>64</b>
<b>Summary</b>	<b>64</b>
<b>Exercise Questions</b>	<b>65</b>
Question 1	65
Question 2	65
<b>Chapter 3: XML Namespaces</b>	<b>67</b>
<hr/>	
<b>Why We Need Namespaces</b>	<b>67</b>
Using Prefixes	69
Why Doesn't XML Just Use These Prefixes?	70
<b>How XML Namespaces Work</b>	<b>72</b>
Default Namespaces	75
Do Different Notations Make Any Difference?	81
Namespaces and Attributes	83
<b>Understanding URIs</b>	<b>86</b>
URLs	86
URNs	87
Why Use URLs for Namespaces, Not URNs?	87
What Do Namespace URIs Really Mean?	88
RDDL	89
<b>When to Use Namespaces</b>	<b>89</b>
<b>Summary</b>	<b>90</b>
<b>Exercise Questions</b>	<b>91</b>
Question 1	91
Question 2	91
Question 3	91

---

<b>Part II: Validation</b>	<b>93</b>
<b>Chapter 4: Document Type Definitions</b>	<b>95</b>
<b>Running the Samples</b>	<b>96</b>
Preparing the Ground	96
The Document Type Declaration	100
<b>Sharing Vocabularies</b>	<b>104</b>
<b>Anatomy of a DTD</b>	<b>105</b>
Element Declarations	105
Attribute Declarations	120
Entities	131
<b>Developing DTDs</b>	<b>141</b>
<b>DTD Limitations</b>	<b>142</b>
DTD Syntax	142
XML Namespaces	143
Data Typing	143
Limited Content Model Descriptions	143
<b>Summary</b>	<b>143</b>
<b>Exercise Questions</b>	<b>144</b>
Question 1	144
Question 2	144
Question 3	144
<b>Chapter 5: XML Schemas</b>	<b>145</b>
<b>Benefits of XML Schemas</b>	<b>146</b>
XML Schemas Use XML Syntax	146
XML Schema Namespace Support	146
XML Schema Data Types	147
XML Schema Content Models	147
<b>Do We Still Need DTDs?</b>	<b>147</b>
<b>XML Schemas</b>	<b>148</b>
The XML Schema Document	148
Running the Samples	148
<schema> Declarations	152
<element> Declarations	155
<complexType> Declarations	165
<group> Declarations	167
Content Models	168
<attribute> Declarations	177
<attributeGroup> Declarations	183

# Contents

---

Creating Elements with Simple Content and Attributes	185
Datatypes	186
<simpleType> Declarations	193
<b>Creating a Schema from Multiple Documents</b>	<b>200</b>
<import> Declarations	200
<include> Declarations	204
<b>Documenting XML Schemas</b>	<b>206</b>
Comments	206
Attributes from Other Namespaces	207
Annotations	208
<b>Summary</b>	<b>209</b>
<b>Exercise Questions</b>	<b>210</b>
Question 1	210
Question 2	210
Question 3	210
<b>Chapter 6: RELAX NG</b>	<b>211</b>
<b>XML and Compact Syntaxes</b>	<b>212</b>
Running the Samples	212
<b>RELAX NG Patterns</b>	<b>213</b>
Element, Attribute, and Text Patterns	213
<b>Combining and Reusing Patterns and Grammars</b>	<b>227</b>
Named Patterns	227
Combining Named Pattern Definitions	230
Schema Modularization Using the include Directive	231
Redefining Included Named Patterns	232
Removing Patterns with the notAllowed Pattern	233
Extensions and Restrictions	234
Nested Grammars	235
<b>Additional RELAX NG Features</b>	<b>236</b>
Namespaces	236
Name-Classes	237
Datatypes	241
List Patterns	243
Comments and Divisions	244
<b>Useful Resources</b>	<b>245</b>
<b>Summary</b>	<b>245</b>
<b>Exercise Questions</b>	<b>245</b>
Question 1	246
Question 2	246

---

<b>Part III: Processing</b>	<b>247</b>
<b>Chapter 7: XPath</b>	<b>249</b>
<b>Ways of Looking at an XML Document</b>	<b>250</b>
<b>Modeling XML Documents</b>	<b>250</b>
<b>Visualizing XPath</b>	<b>251</b>
Understanding Context	252
What Is a Node?	254
XPath 1.0 Types	257
<b>Abbreviated and Unabbreviated Syntax</b>	<b>259</b>
<b>XPath 1.0 Axes</b>	<b>260</b>
Child Axis	260
attribute Axis	262
ancestor Axis	264
ancestor-or-self Axis	265
descendant Axis	265
descendant-or-self Axis	266
following Axis	266
following-sibling Axis	268
namespace Axis	268
parent Axis	271
preceding Axis	271
preceding-sibling Axis	272
self Axis	273
<b>XPath 1.0 Functions</b>	<b>274</b>
Boolean Functions	274
Node-Set Functions	275
Numeric Functions	275
String Functions	276
<b>Predicates</b>	<b>278</b>
<b>Structure of XPath Expressions</b>	<b>278</b>
<b>XPath 2.0</b>	<b>281</b>
Revised XPath Data Model	281
W3C XML Schema Data Types	281
Additional XPath 2.0 Functions	282
XPath 2.0 Features	282
<b>Summary</b>	<b>285</b>
<b>Exercise Questions</b>	<b>285</b>
Question 1	286
Question 2	286

<b>Chapter 8: XSLT</b>	<b>287</b>
<b>What Is XSLT?</b>	<b>287</b>
Restructuring XML	288
Presenting XML Content	288
<b>How an XSLT Processor Works</b>	<b>288</b>
<b>Running the Examples</b>	<b>289</b>
Introducing the Saxon XSLT Processor	289
Installing the Saxon XSLT Processor	290
<b>Procedural versus Declarative Programming</b>	<b>292</b>
Procedural Programming	292
Declarative Programming	292
<b>Foundational XSLT Elements</b>	<b>293</b>
The <xsl:stylesheet> Element	295
The <xsl:template> Element	296
The <xsl:apply-templates> Element	296
<b>Getting Information from the Source Tree</b>	<b>297</b>
The <xsl:value-of> Element	297
The <xsl:copy> Element	299
The <xsl:copy-of> Element	303
<b>Influencing the Output with the &lt;xsl:output&gt; Element</b>	<b>306</b>
<b>Conditional Processing</b>	<b>306</b>
The <xsl:if> Element	306
The <xsl:choose> Element	308
<b>The &lt;xsl:for-each&gt; Element</b>	<b>311</b>
<b>The &lt;xsl:sort&gt; Element</b>	<b>312</b>
<b>XSLT Modes</b>	<b>314</b>
<b>XSLT Variables and Parameters</b>	<b>320</b>
<b>Named Templates and the &lt;xsl:call-template&gt; Element</b>	<b>322</b>
<b>XSLT Functions</b>	<b>323</b>
<b>XSLT 2.0</b>	<b>323</b>
Grouping in Version 2.0	324
Non-XML Input and String Handling	327
Multiple Outputs	330
User-Defined Functions	332
xsl:value-of changes	334
<b>Summary</b>	<b>335</b>
<b>Exercise Questions</b>	<b>335</b>
Question 1	335
Question 2	335

---

<b>Part IV: Databases</b>	<b>337</b>
<b>Chapter 9: XQuery, the XML Query Language</b>	<b>339</b>
<b>Why XQuery?</b>	<b>340</b>
Historical Factors	340
Technical Factors	340
Current Status	341
<b>XQuery Tools</b>	<b>343</b>
Saxon	343
X-Hive.com Online	345
X-Hive Database	346
Tamino Database	346
Microsoft SQL Server 2005	346
Oracle	346
<b>Some XQuery Examples</b>	<b>346</b>
Input Functions	346
Retrieving Nodes	348
Element Constructors	351
The XQuery Prolog	355
Computed Constructors	358
Syntax	359
<b>The XQuery Data Model</b>	<b>360</b>
Shared Data Model with XPath 2.0 and XSLT 2.0	360
Node Kinds	361
Sequences of Node-Sets	361
Document Order	361
Comparing Items and Nodes	361
Types in XQuery	361
Axes in XQuery	361
<b>XQuery Expressions</b>	<b>362</b>
FLWOR Expressions	362
<b>XQuery Functions</b>	<b>368</b>
The concat() Function	369
The count() Function	369
<b>Using Parameters with XQuery</b>	<b>370</b>
<b>User-Defined Functions</b>	<b>371</b>
<b>Looking Ahead</b>	<b>372</b>
Update Functionality	372
Full-Text Search	372
<b>Summary</b>	<b>372</b>

# Contents

---

<b>Exercise Questions</b>	<b>373</b>
Question 1	373
Question 2	373
<b>Chapter 10: XML and Databases</b>	<b>375</b>
<b>The Need for Efficient XML Data Stores</b>	<b>375</b>
The Increasing Amount of XML	376
Comparing XML-Based Data and Relational Data	377
<b>Approaches to Storing XML</b>	<b>378</b>
Storing XML on File Systems	378
Using XML With Conventional Databases	379
Native XML Databases	381
<b>Using Native XML Databases</b>	<b>382</b>
Obtaining and Installing eXist	382
Interacting with eXist	384
<b>XML in Commercial RDBMSs</b>	<b>395</b>
XML Functionality in SQL Server 2000	395
Web Service Support	426
<b>XML in Open Source RDBMS</b>	<b>426</b>
Installing MySQL	426
Adding Information in MySQL	427
Querying MySQL	430
Updating XML in MySQL	435
Usability of XML in MySQL	436
Client-Side XML Support	437
<b>Choosing a Database to Store XML</b>	<b>438</b>
<b>Looking Ahead</b>	<b>438</b>
<b>Summary</b>	<b>438</b>
<b>Exercise Questions</b>	<b>438</b>
Question 1	439
Question 2	439
Question 3	439
<b>Part V: Programming</b>	<b>441</b>
<b>Chapter 11: The XML Document Object Model (DOM)</b>	<b>443</b>
<b>Purpose of the XML DOM</b>	<b>443</b>
Interfaces and Objects	445
<b>The Document Object Model at the W3C</b>	<b>446</b>
XML DOM Implementations	447

---

<b>Two Ways to View DOM Nodes</b>	<b>448</b>
Overview of the XML DOM	448
<b>Tools to Run the Examples</b>	<b>450</b>
Browser Differences	450
<b>The Node Object</b>	<b>458</b>
Properties of the Node Object	458
Methods of the Node Object	461
Loading an XML Document	462
The Effect of Text Nodes	468
The NamedNodeMap Object	471
The NodeList Object	475
The DOMException Object	476
<b>The Document Interface</b>	<b>478</b>
<b>How the XML DOM Is Used in InfoPath 2007</b>	<b>481</b>
<b>Summary</b>	<b>482</b>
<b>Exercise Questions</b>	<b>482</b>
Question 1	482
Question 2	482
<b>Chapter 12: Simple API for XML (SAX)</b>	<b>483</b>
<hr/>	
<b>What Is SAX and Why Was It Invented?</b>	<b>483</b>
A Brief History of SAX	484
Where to Get SAX	485
Setting Up SAX	486
<b>Receiving SAX Events</b>	<b>486</b>
ContentHandler Interface	487
ErrorHandler Interface	504
DTDHandler Interface	509
EntityResolver Interface	510
Features and Properties	510
Extension Interfaces	514
<b>Good SAX and Bad SAX</b>	<b>515</b>
<b>Consumers, Producers, and Filters</b>	<b>516</b>
<b>Other Languages</b>	<b>516</b>
<b>Summary</b>	<b>517</b>
<b>Exercise Questions</b>	<b>518</b>
Question 1	518
Question 2	518

**Part VI: Communication** **519**

**Chapter 13: RSS, Atom, and Content Syndication** **521**

---

<b>Syndication and Meta Data</b>	<b>521</b>
Syndication Systems	522
The Origin of RSS Species	525
RSS-DEV and RSS 1.0	529
UserLand and RSS 2.0	531
Atom	533
<b>Working with News Feeds</b>	<b>536</b>
Newsreaders	536
Data Quality	536
<b>A Simple Aggregator</b>	<b>537</b>
Modeling Feeds	537
Program Flow	540
Implementation	540
Transforming RSS with XSLT	557
<b>Useful Resources</b>	<b>567</b>
<b>Summary</b>	<b>568</b>
<b>Exercise Questions</b>	<b>568</b>
Question 1	569
Question 2	569

**Chapter 14: Web Services** **571**

---

<b>What Is an RPC?</b>	<b>571</b>
<b>RPC Protocols</b>	<b>573</b>
DCOM	573
IIOP	574
Java RMI	575
<b>The New RPC Protocol: Web Services</b>	<b>575</b>
XML-RPC	576
The Network Transport	579
<b>Taking a REST</b>	<b>596</b>
<b>The Web Services Stack</b>	<b>600</b>
SOAP	600
WSDL	601
UDDI	602
Surrounding Specifications	602
<b>Summary</b>	<b>604</b>

---

<b>Exercise Questions</b>	<b>605</b>
Question 1	605
Question 2	605
<b>Chapter 15: SOAP and WSDL</b>	<b>607</b>
<hr/>	
<b>Laying the Groundwork</b>	<b>608</b>
<b>Running Examples in Windows 2003, XP, and 2000</b>	<b>608</b>
<b>The New RPC Protocol: SOAP</b>	<b>608</b>
Just RESTing	612
Basic SOAP Messages	613
More Complex SOAP Interactions	620
<b>Defining Web Services: WSDL</b>	<b>632</b>
<definitions>	633
<types>	633
<messages>	634
<portTypes>	635
<binding>	635
<soap:body>	637
<service>	638
Other Bindings	641
<b>Summary</b>	<b>644</b>
<b>Exercise Questions</b>	<b>644</b>
Question 1	644
Question 2	644
<b>Chapter 16: Ajax</b>	<b>645</b>
<hr/>	
<b>Early Attempts at Asynchronous Updates</b>	<b>645</b>
Microsoft versus Mozilla	647
Cross-Browser Solutions	647
<b>Basic Posting Techniques</b>	<b>649</b>
<b>Transport and Processing on the Server</b>	<b>652</b>
JSON	652
Payment Card Validator	653
The AutoSuggest Box	658
<b>Server-Side Proxies</b>	<b>681</b>
The Currency Converter Proxy	682
<b>Summary</b>	<b>686</b>
<b>Exercise Questions</b>	<b>687</b>
Question 1	687
Question 2	687
Question 3	687

## **Part VII: Display** **689**

### **Chapter 17: Cascading Style Sheets (CSS)** **691**

---

<b>Why Stylesheets?</b>	<b>692</b>
<b>Introducing CSS</b>	<b>693</b>
CSS Properties	694
Inheritance	695
Attaching the Stylesheet to an XML Document	699
Selectors	700
<b>Using CSS for Layout of XML Documents</b>	<b>701</b>
Understanding the Box Model	702
Positioning in CSS	706
<b>Laying Out Tabular Data</b>	<b>719</b>
<b>Links in XML Documents</b>	<b>721</b>
XLink Support in Firefox	721
Forcing Links Using the XHTML Namespace	725
<b>Images in XML Documents</b>	<b>725</b>
<b>Using CSS to Add Content to Documents</b>	<b>726</b>
<b>Attribute Content</b>	<b>729</b>
Attribute Selectors	729
Using Attribute Values in Documents	729
<b>Summary</b>	<b>732</b>
<b>Exercise Questions</b>	<b>732</b>
Question 1	732
Question 2	733
Question 3	733
Question 4	733

### **Chapter 18: XHTML** **735**

---

<b>Separating Style from Content</b>	<b>736</b>
<b>Learning XHTML 1.x</b>	<b>738</b>
Document Type Definitions for XHTML	738
Basic Changes in Writing XHTML	740
Styling XHTML Documents	751
Stricter Documents Make Faster and Lighter Processors	753
XHTML Tools	753
Validating XHTML Documents	754
Validation Pitfalls	756
Mime Types Pitfalls	757

<b>Modularized XHTML</b>	<b>759</b>
Module Implementations	761
XHTML 1.1	761
XHTML Basic	762
<b>What's Next for XHTML</b>	<b>763</b>
<b>Summary</b>	<b>765</b>
<b>Exercise Questions</b>	<b>766</b>
Question 1	766
Question 2	766
<hr/> <b>Chapter 19: Scalable Vector Graphics (SVG)</b>	<hr/> <b>767</b>
<b>What Is SVG?</b>	<b>767</b>
Scalable, Vector, Graphics	768
Putting SVG to Work	769
An SVG Toolkit	769
<b>Getting Started</b>	<b>771</b>
Views and Units	774
The Painter's Model	774
Grouping	776
Transformations	776
Paths	777
Images	780
Text	781
Comments, Annotation, and Metadata	782
Scripting	784
SVG on Your Website	785
<b>Tangram: A Simple Application</b>	<b>786</b>
XHTML Wrapper	787
SVG Shapes	788
Tangram Script	792
<b>Useful Resources</b>	<b>799</b>
<b>Summary</b>	<b>800</b>
<b>Exercise Questions</b>	<b>800</b>
Question 1	800
Question 2	801
<hr/> <b>Chapter 20: XForms</b>	<hr/> <b>803</b>
<b>How XForms Improves on HTML Forms</b>	<b>804</b>
<b>XForms Tools</b>	<b>804</b>
<b>An Illustrative XForms Example</b>	<b>810</b>

# Contents

---

<b>XForms Form Controls</b>	<b>817</b>
The xforms:input Element	817
The xforms:secret Element	818
The xforms:textarea Element	818
The xforms:output Element	818
The xforms:upload Element	819
The xforms:range Element	819
The xforms:trigger Element	820
The xforms:submit Element	820
The xforms:select Element	821
The xforms:select1 Element	822
<b>Constraining XForms Instances</b>	<b>828</b>
The xforms:bind Element	828
W3C XML Schema in XForms	833
Schema or Bind Elements: Which One to Choose?	834
<b>XForms Events</b>	<b>834</b>
<b>The XForms Action Module</b>	<b>835</b>
Developing and Debugging XForms	836
<b>Alternatives to XForms</b>	<b>836</b>
Microsoft InfoPath	836
Adobe LiveCycle	837
HTML Forms	838
<b>Summary</b>	<b>838</b>
<b>Exercise Questions</b>	<b>838</b>
Question 1	838
Question 2	838
<b>Part VIII: Case Study</b>	<b>839</b>
<b>Chapter 21: Case Study: Payment Calculator</b>	<b>841</b>
<hr/>	
<b>Mortgage Calculations</b>	<b>841</b>
<b>What You'll Need</b>	<b>842</b>
<b>Online Loan Calculator</b>	<b>842</b>
Integrating the Calculation Web Service	849
Enhancing the Display with SVG	865
Adding the Frame to the Main Page	868
<b>Summary</b>	<b>872</b>
<b>Chapter 22: Case Study: Payment Calculator — Ruby on Rails</b>	<b>Online</b>
<hr/>	

---

<b>Appendix A: Exercise Solutions</b>	<b>873</b>
<b>Appendix B: XPath Reference</b>	<b>923</b>
<b>Appendix C: XSLT Reference</b>	<b>939</b>
<b>Appendix D: The XML Document Object Model</b>	<b>Online</b>
<b>Appendix E: XML Schema Element and Attribute Reference</b>	<b>Online</b>
<b>Appendix F: XML Schema Datatypes Reference</b>	<b>Online</b>
<b>Appendix G: SAX 2.0.2 Reference</b>	<b>Online</b>
<b>Index</b>	<b>971</b>

