

XML 开发指南

[美] Fabio Arciniegas 著
天宏工作室 译

本书附盘可从本馆主页 <http://lib.szu.edu.cn/>
上由“馆藏检索”该书详细信息后下载,
也可到视听部复制

北京·清华大学出版社

XML 开发指南

Fabio Arciniegas: **XML Developer's Guide**

ISBN 0-07-212648-5

Copyright © 2001 by The McGraw-Hill Companies, Inc.

Original English Language Edition Published by The McGraw-Hill Companies, Inc.

All Rights Reserved.

北京市版权局著作权合同登记号 图字 01-2002-3464 号

本书中文简体字翻译版由美国麦格劳-希尔教育（亚洲）出版公司授权清华大学出版社在中国境内（香港、澳门特别行政区和台湾地区除外）独家出版、发行。

未经出版者书面许可，不得以任何方式复制或抄袭本书的任何部分。

版权所有，翻印必究。

本书贴有 **McGraw-Hill Education** 防伪标签，无标签者不得销售。

书 名：XML 开发指南

译 者：天宏工作室

出 版 者：清华大学出版社（北京清华大学学研大厦，邮编 100084）

<http://www.tup.tsinghua.edu.cn>

<http://www.tup.com.cn>

责任编辑：冯志强

印 刷 者：清华大学印刷厂

发 行 者：新华书店总店北京发行所

开 本：787×960 1/16 印张：41.75 字数：902 千字

版 次：2003 年 1 月第 1 版 2003 年 1 月第 1 次印刷

书 号：ISBN 7-89494-021-6

印 数：0001 ~ 3000

定 价：88.00 元

目 录

第一部分 XML 结构、语法和应用

第一章 XML 基础	3
1.1 简介	4
1.2 标记简介	4
1.3 XML 定义	10
1.3.1 严格的定义	10
1.3.2 面向 Web 的定义	10
1.3.3 面向数据的定义	10
1.4 XML 系列的示意图	10
1.5 XML 语法和主要概念	11
1.5.1 简介	12
1.5.2 XML: 详细介绍	13
1.6 综合应用: 示例 XML 文档	27
1.7 小结	28
第二章 DTD: 特性和技术	29
2.1 简介	30
2.2 DTD 简介	30
2.2.1 DTD 和文档	30
2.2.2 声明元素	30
2.2.3 声明属性	33
2.2.4 声明实体	36
2.3 深入研究 DTD	38
2.3.1 元素类型声明	38
2.3.2 属性列表声明	43
2.4 更深入地讨论实体	51
2.4.1 实体类型	51

2.4.2	外部通用实体	52
2.4.3	内部通用实体	53
2.4.4	未分析实体	53
2.5	参数实体的有用集合	54
2.5.1	条件节	56
2.5.2	内部和外部 DTD 子集	57
2.6	小结	57
第三章	名称空间和 XML 处理简介	58
3.1	简介	59
3.2	名称空间	59
3.2.1	定义和例子	59
3.2.2	例子	60
3.2.3	名称空间声明	61
3.2.4	合格名称	63
3.2.5	结构上的细节	64
3.2.6	一些常见的误解	66
3.3	处理模型	68
3.3.1	范例	68
3.3.2	特性	69
3.4	小结	70

第二部分 XML 的分析和程序化处理

第四章	SAX 和 SAX2 (版本 1.0 和 2.0)	73
4.1	SAX 和 SAX2 背后的基本原理	74
4.2	SAX 1.0/SAX2 的用途	74
4.3	SAX 示例程序	75
4.3.1	定义处理程序	76
4.3.2	在分析程序上注册处理程序并开始分析	77
4.4	SAX2 结构	77
4.5	SAX2 的高级特征	91
4.5.1	筛选器	92
4.5.2	适配器	93
4.6	SAX 应用程序中的常见设计错误	94

4.6.1 使用非内聚性的处理程序	94
4.6.2 过多使用 DefaultHandler	94
4.6.3 数据状态的直接修改（在需要历史数据时）	95
4.7 小结	96
第五章 使用 SAX 和 SAX2 的高级设计	97
5.1 简介	98
5.2 SAX2 应用程序中的生成器模式	99
5.2.1 用途	99
5.2.2 例子	99
5.2.3 结构	102
5.2.4 例子	102
5.2.5 结果	109
5.3 SAX2 应用程序中的命令模式	109
5.3.1 用途	109
5.3.2 例子	109
5.3.3 结构	111
5.3.4 例子	112
5.3.5 结果	116
5.4 责任链	116
5.4.1 用途	116
5.4.2 例子	116
5.4.3 结构	118
5.4.4 例子	119
5.4.5 结果	121
5.5 小结	121
第六章 文档对象模型第 2 级（DOM 2）	122
6.1 简介	123
6.1.1 DOM 的历史	124
6.1.2 DOM 是什么	124
6.2 小型 DOM 应用程序	126
6.2.1 问题	127
6.2.2 解决方案	128
6.3 DOM 结构	130

6.3.1	DOM 示意图	131
6.3.2	核心 DOM 接口	131
6.3.3	综合性的例子	136
6.4	小结	141
第七章	使用 DOM2 的高级设计	142
7.1	简介	143
7.2	DOM2 应用程序中的访问者模式	143
7.2.1	访问者模式的用途	143
7.2.2	例子	143
7.2.3	结构	150
7.2.4	例子	152
7.2.5	结果	153
7.3	DOM2 应用程序中的 Iterator 模式	154
7.3.1	Iterator 的用途	154
7.3.2	例子	154
7.3.3	Iterator 模式的结构	158
7.3.4	例子	159
7.3.5	结果	163
7.4	DOM2 应用程序中的 Mediator 模式	163
7.4.1	用途	163
7.4.2	例子	163
7.4.3	结构	167
7.4.4	例子	168
7.4.5	结果	170
7.5	小结	170

第三部分 与 XML 有关的技术

第八章	XPath——XML 路径语言	173
8.1	简介	174
8.2	XPath——经验方法	174
8.2.1	对样式的初步说明	174
8.2.2	XPath 基础知识	174
8.3	XPath 工具	185

8.3.1 XPath 位置路径试验程序	186
8.3.2 emacs XPath 扩展	187
8.3.3 库	187
8.4 通过程序使用 XPath	188
8.5 XPath 的一些细节	189
8.5.1 位置路径	189
8.5.2 步	190
8.5.3 arc	190
8.5.4 简写的语法	191
8.6 小结	191
第九章 XPointer	192
9.1 简介	193
9.2 XPointer 基础知识	193
9.3 XPointer 模型和语言	194
9.3.1 逻辑路径的简要介绍	194
9.3.2 XPointer 的格式	195
9.4 XPointer 转义	197
9.4.1 XPointer URI 引用转义	197
9.4.2 XPointer XML 转义	198
9.4.3 XPointer 转义	199
9.5 XPointer 对 XPath 的扩展	199
9.5.1 XPointer 向 XPath 添加的新概念	199
9.5.2 XPointer 向 XPath 添加的新函数	201
9.5.3 向 XPath 添加的特性总结	203
9.6 XPointer 工具	204
9.7 小结	205
第十章 XLink	206
10.1 简介	207
10.1.1 XML 数据之间的关系	207
10.2 XLink 元素	209
10.3 使用实例说明扩展的 XLink 和简单的 XLink	211
10.3.1 例子的结构	211
10.3.2 XLink 标记	221

10.3.3	扩展链接	221
10.3.4	简单链接	226
10.3.5	XLink 的表示方法	227
10.3.6	总结: XLink 全局属性和使用规则的分类	228
10.4	抽象 XLink 的例子	229
10.4.1	图表的概念	229
10.4.2	在 XML 中使用 XLink 建立定向图模型	230
10.4.3	在 Java 之后建立图表模型	232
10.4.4	从 XML 图表构造 Java 图表	233
10.5	小结	233
第十一章	XSLT: 转换 XML	234
11.1	简介	235
11.2	专业概述	235
11.2.1	基础知识	235
11.2.2	简单的 XSLT 转换	236
11.3	关于转换过程的更多内容	239
11.3.1	在节点上操作	239
11.3.2	直观地了解转换流程	240
11.4	XSLT 处理程序	241
11.4.1	安装和调用 Xalan	242
11.4.2	安装和激活 XT	242
11.5	完整的 XSLT 语言	243
11.5.1	源文档	243
11.5.2	插入新元素和属性	245
11.5.3	插入文本	248
11.5.4	插入生成的文本	249
11.5.5	复制	250
11.5.6	编号	250
11.5.7	条件	252
11.5.8	循环	255
11.5.9	排序	255
11.5.10	变量	258
11.5.11	命名模板	260
11.5.12	向模板传递参数	261

11.6 小结	264
第十二章 XSLT: 高级用法和技巧	265
12.1 简介	266
12.2 从结构到显示: 功能需求	266
12.2.1 需求	266
12.2.2 建立信息模型	266
12.2.3 转换到 XHTML	270
12.2.4 有关转换到 PDF 的说明	276
12.3 非文字表示: 使用 SVG 生成图形	277
12.3.1 图形表示	277
12.3.2 建立信息模型	278
12.3.3 将数据显示为 SVG	280
12.3.4 将统计数据转换为多幅 SVG 图形	282
12.4 样式表重用技术	285
12.5 扩展 XSLT	286
12.5.1 扩展 XSLT 的时机	286
12.5.2 非递归方案	286
12.5.3 实现 times 结构	287
12.5.4 有关其他扩展类型的说明	290
12.5.5 其他扩展	290
12.6 简明的 XSLT 参考	290
12.7 小结	295
第十三章 XML 模式	296
13.1 简介	297
13.2 XML 模式: 专业方法	297
13.2.1 类型	298
13.2.2 侧面	304
13.2.3 匿名类型	309
13.2.4 内容模型	310
13.2.5 元素和属性重用: 命名组	312
13.2.6 注释	313
13.3 小结	313

第十四章 高级数据建模和 XML 模式	315
14.1 简介	316
14.2 XML 模式的高级概念	316
14.2.1 派生和其他高级类型关系	316
14.2.2 惟一性	326
14.2.3 模块化和重用	327
14.3 XML 模式与 DTD 的比较 (Rick Jelliffe)	327
14.4 小结	330

第四部分 主要的 XML 应用

第十五章 无线 XML: WAP、VoiceXML 及其他	335
15.1 简介	336
15.2 WAP	336
15.2.1 WAP 的基本原理	337
15.2.2 WAP 模型	337
15.2.3 WAP 的结构	339
15.2.4 WAE	339
15.3 WML	340
15.3.1 设置 WML 开发环境	341
15.3.2 WML 的用法	342
15.3.3 WML 的完整介绍	344
15.3.4 结构	357
15.4 WMLScript	362
15.4.1 WMLScript 结构	362
15.4.2 VoiceXML	369
15.5 小结	370
第十六章 XML 和数据库	371
16.1 简介	372
16.2 XML-数据库关系	372
16.2.1 文档类型	372
16.2.2 耦合程度	375
16.3 特定词汇的自定义类	376
16.4 用于独立数据库的 API	380

16.5	中间件工具	386
16.5.1	模板与映射	387
16.5.2	xml-dbms	388
16.5.3	综合应用：完整的例子	390
16.6	元数据转换	395
16.7	内容管理系统	396
16.8	XML（数据库应用程序）服务器	396
16.9	产品列表	399
16.10	小结	401
第十七章	XML 服务器到服务器：XML-RPC 和 B2B	402
17.1	简介	403
17.2	XML-RPC	403
17.3	远程过程调用基础	403
17.3.1	调用的形式	404
17.3.2	响应的形式	405
17.3.3	错误的形式	406
17.3.4	测试 XML-RPC 服务	407
17.3.5	XML-RPC 的优缺点	407
17.3.6	XML-RPC 编程的专业介绍	409
17.3.7	深入了解 XML-RPC 规范	413
17.4	使用 XML-RPC 的 B2B 应用程序：DVD 价格比较	420
17.4.1	结构	421
17.4.2	DVDTitles 服务	422
17.4.3	DVDTitles 服务器	425
17.4.4	DVDPage 客户	426
17.4.5	DVDPrices 服务	427
17.5	小结	431

第五部分 综合实例

第十八章	表示 XML：CSS2、XHTML、SVG 和 SMIL	435
18.1	简介	436
18.2	XML + CSS	436
18.2.1	XML + CSS 简介	436
18.2.2	语法和机制	443

18.2.3	CSS 可视方框模型	454
18.2.4	完整的例子	455
18.3	小结	462
第十九章	错误跟踪应用程序	474
19.1	简介	475
19.2	需求	475
19.2.1	XMLBugTrack 的需求 (由 req2txt.xsl 生成)	475
19.2.2	XMLBugTrack 的需求 (由 req2xhtml.xsl 生成)	477
19.2.3	XMLBugTrack 的需求 (原始的 XML)	478
19.3	分析与设计	482
19.3.1	结构图	482
19.3.2	设计	483
19.3.3	实现	488
19.3.4	从数据库到 XML 的连接和 XSLT 处理	489
19.4	小结	507
第二十章	知识管理应用程序	508
20.1	简介	509
20.2	什么是知识管理?	509
20.3	需求	510
20.4	底层的 DTD	515
20.4.1	注释 DTD	515
20.4.2	字典	525
20.5	注释文档的编辑器	528
20.6	创建表示方法的 SMIL/HTML	531
20.7	小结	541
 第六部分 附录		
附录 A	XML 语法参考	545
附录 B	XML 1.0 规范	548
附录 C	主要的表示型 DTD	604
附录 D	UML 快速参考	641
附录 E	重要的自定义 DTD	644

第一部分

XML 结构、语法和应用

第一章

XML 基础

1.1 简介

本章将介绍 XML 作为一种语言的基础知识。对于某些读者而言，阅读本章非常重要，并且应该在继续阅读本书其余内容之前掌握这一章的内容。不过，拥有丰富经验或者已经对 XML 有较深入了解的读者可以略过本章，并在必要时返回这一章。

本章分为两部分：一部分主要介绍历史和概念上的背景知识，而另一部分纯粹是技术性的。第一部分介绍 XML 的历史和定义，而第二部分直接介绍 XML 作为一种语言的语法和语义。

1.2 标记简介

我们经常会听到 XML 的如下定义：

- XML 是 SGML (Standard Generalized Markup Language, 标准通用标记语言) 的一个子集。
- XML 是一种通用标记语言。
- XML 是创建标记语言的一种元语言。

引用时髦的词语与真正掌握一项技术之间的一个重要区别就是正确理解技术知识和术语。因此，让我们首先检查这些定义的来源以及它们如何真正应用于 XML。

什么是标记

标记这个术语适用于向文档内容添加的任何一组代码或标签（以便指出它的含义或显示方式）。

预标记和特定的标记语言

如果没有标记，那么文档将被局限于只显示内容数据。文档中没有明确的结构，因此必须在文档外部对应用于数据的任何特殊含义或处理过程进行编码。

标记允许在文档中包括元数据。不过，正如无数专用标记机制所表明的，这并不意味着文档在语义上更丰富，更容易使用。

专用标记机制在很早以前就出现了，每一种机制都对它们定义的词汇采用不同的约定和含义。考虑 RTF (Rich Text Format, 多信息文本格式) 的例子，如图 1.1 中的 Windows WordPad 所示。