



今日电子

美国计算机“宝典”丛书

XML Bible 2nd Edition

丛书
累计印数
64万册

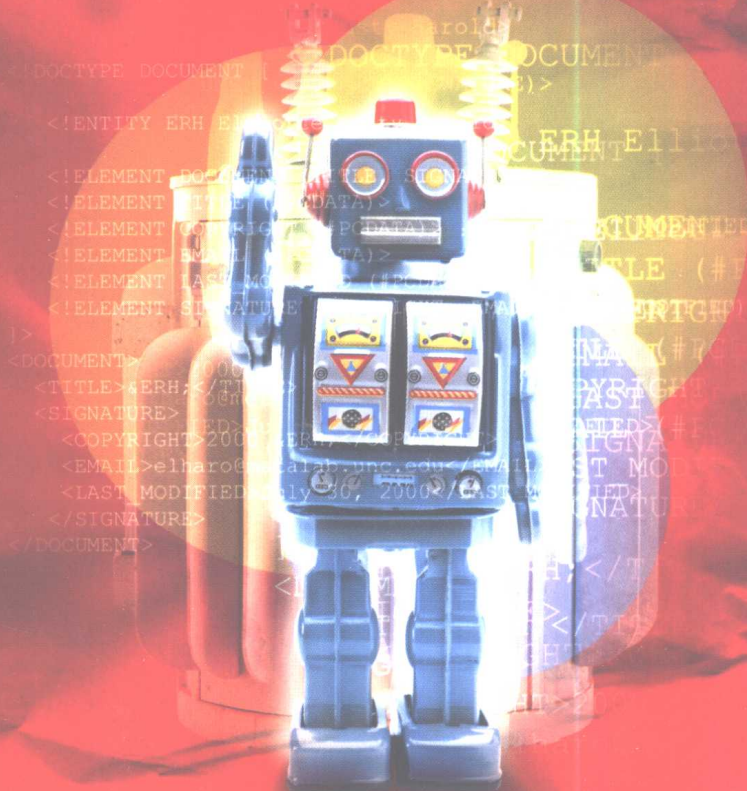
100%

内容丰富、权威

掌握XML的基础知识,
包括元素、标记、属
性、DTD和名字空间

利用强大的CSS和XSL
来格式化XML文档

使用XLink、XPointer、
模式、SVG和XHTML将
XML的功能发挥到极致



XML

[美] Elliott Rusty Harold 著

马云钟萍等译

陈孟建 审校

随书
附带的
CD-ROM
光盘包含
XML示例代码、
XML的创作工具
和W3C标准

宝典

(第二版)



电子工业出版社

Publishing House of Electronics Industry
www.phei.com.cn

美国计算机“宝典”丛书

XML 宝典

(第二版)

XML Bible

2nd Edition

[美] Elliotte Rusty Harold 著

马 云 钟 萍 等译

陈孟建 审校

電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

可扩展标记语言 (XML) 是由万维网联盟 (W3C) 发布的用于网页设计和数据交换的新技术, 具有良好的发展前景。本书是一本全面介绍 XML 的优秀书籍, 共分为五部分。书中首先介绍了 XML 的基本概念和基本语法, 接着讲解了如何利用文档类型定义 (DTD) 描述和验证 XML 文档, 如何声明文档的各个组成部分以及在文档中嵌入非 XML 的数据。利用本书, 读者能够掌握如何使用样式表 (CSS 和 XSL) 处理文档的显示效果, 并将文档转化为 HTML, 用于在浏览器中显示。书中还介绍了 XML 的相关技术 XLink 和 XPointer, 这些技术为形成文档之间和文档内部的链接提供了最有效的工具。最后, 本书向读者展示了 XML 在不同领域中的应用实例, 并在附录中给出了 XML 参考资料以及 XML 1.0 规范第二版的完整内容。

本书提供了大量的 XML 应用实例, 随书所带的光盘上涵盖了书中的所有实例以及与 XML 相关的浏览器和解析器, 这些都为读者学习 XML 提供了极大的帮助。本书既可以作为学习 XML 的入门教程, 也可以是软件开发人员、网页制作者必不可少的权威参考手册。

Hungry Minds Copyright ©2002 by Publishing House of Electronics Industry. Original English language edition copyright ©2001 by Hungry Minds, Inc. All rights reserved including the right of reproduction in whole or in part in any form. This translation published by arrangement with Hungry Minds, Inc. 本书中文简体专有翻译出版权由美国 Hungry Minds, Inc. 授予电子工业出版社及其所属今日电子杂志社。未经许可, 不得以任何手段和形式复制或抄袭本书内容。该专有出版权受法律保护, 侵权必究。

图书在版编目 (CIP) 数据

XML 宝典 (第二版) / (美) 哈罗德 (Harold, E. R.) 著; 马云等译. - 北京: 电子工业出版社, 2002.1
(美国计算机“宝典”丛书)

书名原文: XML Bible 2nd Edition

ISBN 7-5053-7434-6

I. X... II. ①哈... ②马... III. 可扩展语言, XML-程序设计 IV. TP312

中国版本图书馆 CIP 数据核字 (2001) 第 097828 号

丛 书 名: 美国计算机“宝典”丛书

书 名: XML 宝典 (第二版)

原 书 名: XML Bible 2nd Edition

著 者: [美] Elliotte Rusty Harold

译 者: 马 云 钟 萍 等

审 校 者: 陈孟建

责任编辑: 冯小贝

排版制作: 今日电子公司制作部

印 刷 者: 北京东光印刷厂

出版发行: 电子工业出版社 www.phei.com.cn

北京市海淀区万寿路 173 信箱 邮编: 100036

经 销: 各地新华书店

开 本: 787 × 1092 1/16 印张: 52.75 字数: 1350 千字

版 次: 2002 年 1 月第 1 版 2002 年 1 月第 1 次印刷

书 号: ISBN 7-5053-7434-6
TP · 4286

定 价: 79.00 元 (含光盘一张)

著作权合同登记号 图字: 01-2001-2154

凡购买电子工业出版社的图书, 如有缺页、倒页、脱页、所附磁盘或光盘有问题者, 请向购买书店调换。若书店售缺, 请与本社发行部联系调换。电话: 88211980 68279077

出版说明

21 世纪是一个崭新的世纪，是催人奋进的世纪。在新世纪的第一乐章中，我们热忱地向广大读者、IT 人士推介这套全新改版的美国计算机“宝典”丛书。

丛书的出版宗旨

本着提高广大读者计算机专业技能的宗旨，我社从美国 Hungry Minds(原IDG Books Worldwide)公司引入了这套“宝典”丛书。该丛书在世界各地 51 个国家被译为 31 种文字，拥有几百万的读者。在我国自 1994 年引入这套丛书以来，截至 2001 年中期累计销量已近 60 余万册。得到了广大读者的认同，成为电子工业出版社的著名品牌之一。

丛书的涉及范围

“宝典”丛书的涉及范围甚广，既包括众多的流行软件、编程语言、图形图像，也包括数据库、网络等高端技术等方面的书籍。对于某些软件，我们还进行了本地化处理，按相应的中文版软件进行了调整，进一步贴近中国读者的需求。

每一本“宝典”共同贯彻的一项宗旨就是，对相应主题的介绍都非常全面、系统。使该软件或系统能做到的，读者通过本书的学习也能做到。

丛书的创作队伍

“宝典”丛书的作者都是某个计算机专业领域的专家、教授，有些还是某软件的特约测试者。比如 Deke McClelland、Alan Simpson 和 Ellen Finkelstein 等知名畅销计算机图书作家，在相关领域都具有很高的声望。他们拥有丰富的实践经验，所介绍的内容都是在工作中得到千锤百炼，具有一定的权威性。在他们所撰写的书籍当中，会介绍一些技巧，同时也会为读者提出某些忠告，以免犯同样的错误。

在中文版“宝典”中我们也本着同样的原则，所选的译者均经过严格筛选，他们大都是来自于高等院校的教授、学者，计算机领域的高手，不但具有高深的专业知识，同时也具备英语方面的深厚底蕴。我们的编辑队伍，同样是来自于计算机专业的高素质人才。通过这种严格的层层把关，我们相信最终奉献给读者的将是一部部精品。

丛书的新特性

进入新的世纪，“宝典”将以全新的面貌呈现在广大读者面前。无论是版式、用纸还是印刷质量，相关人员都颇费一番苦心，进行了很大改善。同时我们对于丛书的选题也进行了调整，使其更适合我国的计算机发展水平。对于原书中某些不适合中国国情以及过于调侃的内容进行了删减。我们将秉承“宝典”丛书一贯的“权威、全面、精益求精”的风格，力争使每一本书成为您探索计算机领域奥秘的“宝典”。

译 者 序

读者应该已能感受到，XML在网络上的应用已初露矛头。HTML本身的缺陷以及所带来的不便已经阻碍了网络技术的发展，使人们受到越来越多的限制，而XML的简单性、严格性以及良好的数据传输性，将成为下一代网络技术发展的核心。XML及其相关技术即将带来网络技术的又一次革命。

本书将带领读者进入全新的XML世界，使读者了解XML的方方面面。本书介绍了如何编写XML文档，如何定义和声明XML的元素、属性、实体等各个部分，如何验证文档的有效性，以及如何通过CSS和XSL设计文档的外观。书中还介绍了XML的补充技术，使用XLink和XPointer链接不同的文档和文档内部的不同内容；列举了几种不同的XML应用，从XHTML、无线标记语言，到可伸缩的矢量图形、频道定义格式。XML已经涉及了许多应用领域，甚至还关系到人们生活的方方面面。

本书按照由浅到深的方式全面介绍了XML，并以大量的实例文档帮助读者理解XML、学习XML以及应用XML。随书附带的光盘上，还包含了本书的源代码、示例、本书英文版的PDF文件以及其他的工具和实用程序。相信这本书会成为读者学习XML的最权威、最有效的参考资料。

马云、钟萍、叶喜涛、王仑、赵良叶、尹岩青、袁民、卢朝晖、刘淑珍、卢凌云、王少华、施波、齐建峰、邸伟、李明、耿蕊、王勇、王丽、陆迪泉、于爽、易军凯翻译了本书，全书由陈孟建审校。在本书的翻译和整理过程中，韩军、王海涛、尹燕、张力、韩云萍做了大量工作并提出了宝贵意见，在此对他们表示衷心的感谢。

目 录

前言	1
第 1 部分 XML 基础知识	7
第 1 章 XML 概述	8
1.1 XML 是什么	8
1.2 为什么开发人员如此青睐 XML	10
1.3 XML 文档的生命周期	12
1.4 相关的技术	13
1.5 小结	17
第 2 章 XML 应用程序	18
2.1 XML 应用程序	18
2.2 描述 XML 的 XML	39
2.3 未公开的 XML 应用	41
2.4 小结	44
第 3 章 第一个 XML 文档	45
3.1 Hello XML	45
3.2 研究这个简单的 XML 文档	47
3.3 给 XML 标记指定含义	48
3.4 为 XML 文档编写一个样式表	48
3.5 将一个样式表附加到一个 XML 文档上	49
3.6 小结	50
第 4 章 组织数据	51
4.1 检查数据	51
4.2 将数据组织成 XML 格式	55
4.3 XML 格式的优点	66
4.4 为文档显示准备样式表	67
4.5 小结	77
第 5 章 属性、空标记和 XSL	78
5.1 属性	78

5.2 属性与元素	82
5.3 空元素和空元素标记	87
5.4 XSL	88
5.5 小结	107
第 6 章 形式规范	109
6.1 形式规范的规则	109
6.2 XML 文档	110
6.3 XML 内的文本	112
6.4 元素和标记	112
6.5 属性	115
6.6 实体引用	117
6.7 注释	118
6.8 处理指令	120
6.9 CDATA 段	121
6.10 形式规范的 HTML	121
6.11 小结	131
第 7 章 外国语言和非罗马字符的文本	132
7.1 Web 上非罗马字符的文本	132
7.2 文字、字符集、字体和字形	136
7.3 遗留的字符集	139
7.4 Unicode 字符集	145
7.5 使用 Unicode 编写 XML 文档	149
7.6 小结	152
第 2 部分 文档类型定义	153
第 8 章 DTD 和有效性	154
8.1 文档类型定义	154
8.2 元素声明	155
8.3 DTD 文件	156
8.4 文档类型声明	156
8.5 根据 DTD 进行验证	161
8.6 小结	164
第 9 章 元素声明	166
9.1 分析文档	166
9.2 ANY 内容模型	170
9.3 #PCDATA 内容模型	171

9.4 子元素	173
9.5 序列	174
9.6 一个或多个子元素	175
9.7 零个或多个子元素	175
9.8 零个或一个子元素	176
9.9 用圆括号进行分组	178
9.10 选择	180
9.11 混合内容	181
9.12 空元素	181
9.13 DTD 中的注释	182
9.14 小结	187
第 10 章 实体声明	188
10.1 什么是实体	188
10.2 内部常规实体	189
10.3 外部常规实体	192
10.4 内部参数实体	195
10.5 外部参数实体	197
10.6 从多个文档片段建立完整的文档	201
10.7 小结	209
第 11 章 属性声明	210
11.1 属性的定义	210
11.2 在 DTD 中声明属性	210
11.3 声明多个属性	211
11.4 指定属性的默认值	212
11.5 属性类型	213
11.6 预定义的属性	218
11.7 基于属性的棒球统计数据的 DTD	223
11.8 小结	229
第 12 章 未解析的实体、注记和非 XML 数据	230
12.1 注记	230
12.2 未解析的实体	232
12.3 处理指令	236
12.4 DTD 中的条件部分	238
12.5 小结	239
第 13 章 名字空间	240
13.1 名字空间的需求	240
13.2 名字空间的语法	241

13.3 名字空间和有效性	253
13.4 小结	253
第 3 部分 样式语言	255
第 14 章 CSS 样式表	256
14.1 什么是级联样式表	256
14.2 CSS 中的注释	260
14.3 选择元素	261
14.4 继承	268
14.5 级联	270
14.6 对于不同媒体的不同规则	271
14.7 导入样式表	272
14.8 样式表的字符集	272
14.9 小结	273
第 15 章 CSS 布局	274
15.1 CSS 中的单位	274
15.2 显示属性	280
15.3 框属性	289
15.4 大小属性	297
15.5 定位	301
15.6 格式化页面	306
15.7 小结	309
第 16 章 CSS 文本样式	310
16.1 字体属性	310
16.2 color 属性	319
16.3 文本属性	320
16.4 背景属性	328
16.5 visibility 属性	335
16.6 cursor 属性	336
16.7 content 属性	337
16.8 有声样式表	343
16.9 小结	347
第 17 章 XSL 转换	349
17.1 什么是 XSL	349
17.2 XSL 转换的概述	350
17.3 XSL 模板	358

17.4 使用 <code>xsl:value-of</code> 计算节点值	361
17.5 使用 <code>xsl:for-each</code> 处理多个元素	362
17.6 匹配节点的模式	363
17.7 选择节点的 XPath 表达式	373
17.8 默认的模板规则	385
17.9 确定输出要包含的内容	386
17.10 使用 <code>xsl:copy</code> 复制上下文节点	392
17.11 使用 <code>xsl:number</code> 对节点计数	393
17.12 对输出元素排序	398
17.13 <code>mode</code> 属性	400
17.14 使用 <code>xsl:variable</code> 定义常量	402
17.15 命名的模板	403
17.16 向模板传递参数	404
17.17 删除和保留空格	405
17.18 进行选择	406
17.19 合并多个样式表	407
17.20 输出方式	409
17.21 小结	412
第 18 章 XSL 格式化对象	414
18.1 格式化对象及其属性	414
18.2 页面布局	422
18.3 内容	433
18.4 前导线和水线	436
18.5 图形	438
18.6 链接	442
18.7 列表	443
18.8 表格	446
18.9 行内对象	451
18.10 脚注	451
18.11 浮动内容	452
18.12 格式化属性	452
18.13 小结	466
第 4 部分 补充技术	469
第 19 章 XLink	470
19.1 XLink 和 HTML 链接的对比	470
19.2 链接元素	471
19.3 远程资源的描述	473

19.4 链接行为	474
19.5 扩展链接	477
19.6 扩展链接语法	478
19.7 弧	481
19.8 行外链接	486
19.9 小结	490
第 20 章 XPointer	492
20.1 为什么使用 XPointer	492
20.2 XPointer 的例子	493
20.3 具体例子	495
20.4 定位路径、步骤和集合	497
20.5 根节点	498
20.6 轴	499
20.7 节点测试	502
20.8 谓词	504
20.9 返回节点集合的函数	506
20.10 点	508
20.11 范围	509
20.12 子元素序列	511
20.13 小结	512
第 21 章 资源描述框架	513
21.1 什么是 RDF	513
21.2 RDF 语句	513
21.3 基本 RDF 语法	515
21.4 缩写的 RDF 语法	520
21.5 容器	521
21.6 RDF 模式	528
21.7 小结	529
第 5 部分 XML 应用程序	531
第 22 章 XHTML	532
22.1 验证 HTML 的原因	532
22.2 介绍 XHTML	533
22.3 XHTML 中的新特性	562
22.4 小结	571

第 23 章 无线标记语言	572
23.1 什么是 WML	572
23.2 Hello WML	573
23.3 基本的文本标记	577
23.4 卡片和链接	581
23.5 选择	586
23.6 选项菜单	587
23.7 事件	589
23.8 首部	591
23.9 变量	592
23.10 服务器	597
23.11 小结	599
第 24 章 模式	601
24.1 DTD 的不足之处	601
24.2 模式是什么	602
24.3 W3C 的 XML 模式语言	603
24.4 Hello 模式	604
24.5 复杂类型	607
24.6 分组	615
24.7 简单类型	618
24.8 派生类型	623
24.9 空元素	628
24.10 属性	629
24.11 名字空间	631
24.12 注释	636
24.13 小结	637
第 25 章 可伸缩的矢量图形	640
25.1 SVG 是什么	640
25.2 一个简单的 SVG 文档	641
25.3 在网页中嵌入 SVG 图片	644
25.4 简单形状	646
25.5 路径	652
25.6 文本	657
25.7 位图化的图形	661
25.8 坐标系和视窗	663
25.9 将形状分组	667
25.10 引用形状	668
25.11 转换	670

25.12 链接	675
25.13 元数据	676
25.14 SVG 编辑器	678
25.15 小结	678
第 26 章 矢量标记语言	680
26.1 什么是 VML	680
26.2 用键盘绘图	682
26.3 用 CSS 定位 VML 形状	687
26.4 Microsoft Office 中的 VML	692
26.5 小结	697
第 27 章 频道定义格式	698
27.1 什么是频道定义格式	698
27.2 创建频道	698
27.3 描述频道	701
27.4 更新时间表	705
27.5 提前缓存和 Web 爬行	707
27.6 读者访问日志	708
27.7 BASE 属性	710
27.8 LASTMOD 属性	711
27.9 USAGE 元素	712
27.10 小结	718
第 28 章 设计新的 XML 应用程序	719
28.1 数据的组织	719
28.2 个人的 DTD	724
28.3 家庭的 DTD	728
28.4 信息源的 DTD	729
28.5 家谱树 DTD	730
28.6 为家族树设计一个样式表	736
28.7 小结	741
附录 A CD-ROM 上的内容	742
附录 B XML 参考资料	745
附录 C XML 1.0 规范第二版	790

前言

欢迎阅读《XML 宝典》的第二版。两年前本书推出了第一版，那时的 XML 是一种使用范围较小但却逐渐扩大的新兴技术。近两年来 XML 飞速地增长，不再需要证明它是一种很好的语言规范。事实上，开发人员的问题从“为什么要使用 XML？”转化为“为什么不使用 XML？”在股票交易软件和图形设计软件中，XML 已经成为可选的数据格式。目前使用 XML 的新程序比不使用的要多。因此，推出一部介绍 XML 是什么以及如何使用的计算机书籍应该是很有必要的。

《XML 宝典》是一部带领读者进入引人入胜、快速发展的 XML 世界的入门书籍。通过本书，读者将学会如何编写 XML 文档，如何使用样式表将这些文档转化为 HTML，以便于老式的浏览器查看。读者还会学到如何使用文档类型定义 (DTD) 来描述文档和验证文档。本书将会介绍很多领域的 XML 应用程序，从经济到矢量图形再到家谱的研究。读者将学会如何将 XML 的优势应用于自己的项目、程序和 Web 站点中。

本书的使用者

与市场上其他很多的 XML 书籍不同，《XML 宝典》从网页制作者的角度而不是从软件开发人员的角度介绍 XML。本书不准备花很多时间讨论 BNF 语法或者解析元素的树形结构，而是着重讲解如何使用 XML 和现有的工具开发漂亮、易用、易于维护的 Web 站点，从而吸引人们经常来访。

本书直接针对 Web 开发人员，并且假设读者希望使用 XML 创建 Web 站点，而这个站点很难用原始的 HTML 实现。读者会发现将样式表和自由软件工具结合以后，XML 可以轻松地实现许多功能，而以前这些功能需要使用昂贵的自定义软件，或者需要大量的编程语言（如 Perl）知识。本书使用的软件只需要几分钟的下载时间，而且不需要特殊的编程技巧。

本书第二版新增的内容

本书第二版是从头到尾重新编写的。在保持本书第一版中普遍受欢迎的基本特征和框架以外，第二版的内容更加紧凑，并且尽量解决第一版存在的不恰当之处。例如，原来较大的例子现在变得更小、更易于理解；更正了一些内容错误和印刷错误；最重要的是所编写的正文内容完全符合 2001 年 XML 的最新标准。这一段时间有很多技术发生了巨大变化，1999 年讨论的很多工具（例如 XSLT、XSL-FO、XHTML、XLink、XPointer 和名字空间等），现在已经成为了未来 XML 技术的基础。因此现在可以对这些技术提供更全面的介绍，而不会像第一版那样进行试探性的介绍。

世界是在不断变化的。在本书第一版出版后的两年内，XML 的新技术以惊人的速度发布，本书第二版也对它们进行了介绍，尽管其中的细节还在不断地变换。第二版有几章内容是全新的，并介绍了边缘领域的应用程序，包括：

- ◆ 可扩展超文本标记语言 (Extensible Hypertext Markup Language, 缩写为 XHTML)
- ◆ 可伸缩的矢量图形 (Scalable Vector Graphics, 缩写为 SVG)
- ◆ 模式 (Schema)
- ◆ 无线标记语言 (Wireless Markup Language, 缩写为 WML)

穿插在第一版已有章节中的新内容比新章节更重要。在第二版中,虽然笔者尽量编写得简洁清楚,但最后本书还是比前一版增加了200页,这200页是分散在本书各处的新内容。如果读者喜欢本书的第一版,就会更喜欢第二版,它的方方面面都比第一版更好、更全面和更准确。如果读者不喜欢第一版,也许会发现第二版将满足自己的需求。

读者需要知道的内容

XML是建立在 Internet 和 Web 底层结构之上的,读者需要知道如何使用 ftp 传输文件、发送电子邮件、在某个浏览器中载入 URL,并且还要具备一定的 HTML 知识,例如 Netscape 1.1 所支持的 HTML 的级别。另外,在讨论还没有广泛使用的 HTML 新特性时,如级联样式表 (Cascading Style Sheet, 缩写为 CSS),所讨论的内容都达到了一定深度。

更具体的来讲,要求读者具备以下条件:

- ◆ 使用文本编辑器编写基本的 HTML 页面,包括超链接、图像和文本。
- ◆ 将页面放到 Web 服务器上。

另外,不要求读者具备的知识和技能是:

- ◆ 熟悉 SGML。实际上,SGML 仅仅出现在本书的前言部分。XML 应该比 SGML 更简单和更普遍,因此读者不需要首先学习 SGML。
- ◆ 读者是否为 Java、Perl、C 或者其他编程语言的程序员。XML 是一种标记语言,而不是一种程序设计语言。编写 XML 文档不要求具备编程经验。

读者将要学到的内容

本书的主要目的是教读者为 Web 编写 XML 文档。XML 类似于 HTML (而不是 SGML),可以一步步深入学习。读者掌握了一些内容后就能够进行一些设计,掌握了更多内容就能够进行更多的设计和开发。因此本书的内容从头到尾是很连贯的,读者最好按顺序阅读,在这个过程中将学到以下内容:

- ◆ 如何编写 XML 文档并对外发布。
- ◆ 为什么 XML 文档的语法标记比 HTML 的等价标记更容易开发和维护。
- ◆ 如何按照任何人都可以阅读的格式将 XML 文档发送到 Web 服务器上。
- ◆ 如何保证 XML 的形式规范。
- ◆ 如何在文档中使用其他国家的字符,例如 Ж 和 Å。
- ◆ 如何根据 DTD 和模式来验证文档的正确性。
- ◆ 如何使用实体,从而用较小的模块建立较大的文档。
- ◆ 如何使用属性来描述数据。
- ◆ 如何在文档中嵌入非 XML 的数据。
- ◆ 如何使用名字空间来合并不同的 XML 符号集。
- ◆ 如何使用 CSS 和 XSL 样式表格式化文档。

- ◆ 如何使用 XLink 和 XPointer 链接文档。
- ◆ 如何使用 RDF 为网页编写元数据。

在本书的最后一部分，将介绍几个在实际应用程序中使用 XML 的例子，其中包括：

- ◆ Web 站点设计
- ◆ 模式
- ◆ 推送技术
- ◆ 矢量图形
- ◆ 家谱研究

本书的组织结构

本书分为五个部分：

- 第 1 部分：XML 基础知识
- 第 2 部分：文档类型定义
- 第 3 部分：样式语言
- 第 4 部分：补充技术
- 第 5 部分：XML 应用程序

读者阅读完本书后，就可以使用 XML 创建引人注目的网页。下面分别说明这五个部分。

第 1 部分：XML 基础知识

第 1 部分首先介绍 XML 的历史背景和相关理论，以及 XML 要达到的目标。这一部分展示了如何将 XML 的不同部分组合到一起，实现文档的创建和发送。读者将会看到几个引人注目的 XML 应用程序的例子，这些例子会让读者对 XML 广泛的适用性留下深刻的印象。其中包括可伸缩的矢量图形 (SVG)、资源描述框架 (RDF)、数学标记语言 (MathML)、可扩展的表单描述语言 (XFDL) 等等。然后通过例子向读者介绍如何使用自定义的标记编写 XML 文档，如何在一个文本编辑器中编辑文档，为文档引入样式表、将它们载入类似 Internet Explorer 5.0 或者 Mozilla 的浏览器中。读者还将学习如何编写非英语语言的 XML 文档，例如中文、希伯来语和俄语等。

第 2 部分：文档类型定义

第 2 部分重点介绍文档类型定义 (DTD)。DTD 可以详细说明允许或者不允许哪些元素出现在 XML 文档中，以及这些元素的确切内容和结构。一个进行验证的解析器可以读取文档，并将文档与其 DTD 进行比较，然后报告所发现的任何错误。DTD 使文档创建者能够确保其工作满足所需的任何标准。

在第 2 部分中，读者将学习如何将 DTD 附加到一个文档上，如何使用 DTD 验证文档的正确性，如何编写自己的 DTD 来解决遇到的问题。读者将会掌握定义元素、属性、实体和符号的语法，还可以了解到如何使用实体声明和实体引用，从而根据多个无关的片段确立文档和 DTD。这种功能可以将很长的、难于理解的文档划分为相关的模块和组件，从而使文档得到简化。读者还将学习

如何将其他格式的数据集成到 XML 文档中,例如原始的文本和 GIF 图像文件。这个部分还讲解了如何使用名字空间将一个文档中不同的 XML 符号集混合在一起。

第3部分: 样式语言

第3部分介绍了有关样式表的所有知识。与 HTML 不同,XML 标记仅仅规定文档中的内容而不说明内容的样式;而有关 XML 文档在打印时、在浏览器中查看时或者以其他方式显示时的外观信息则保存在样式表中。可以将不同的样式表应用于相同的文档中,例如可以使用一个样式表为文档的打印设置小字体,而使用另外一个样式表为屏幕显示指定大字体,第三种使用非常大的字体将文档投影到墙上。可以通过选择不同的样式表改变 XML 文档的外观而不涉及文档内容本身。

第3部分详细描述了当前应用最广泛的两种样式表:级联样式表(CSS)和可扩展样式表(XSL)。CSS是一种简单的样式表语言,最初用于 HTML。CSS 存在两个版本:CSS1 和 CSS2。CSS1 提供有关字体、颜色、位置和文本属性的基本信息,该版本已经得到了目前解析 HTML 和 XML 的浏览器的广泛支持。CSS2 是一个较新的标准,其中增加了对有声样式表、用户界面样式表、国际化和双向文本等等的支持。

相比之下,XSL 是一种更复杂、功能也更强大的样式语言,它可以将样式应用于元素的内容,也可以重新组织元素,并增加模板文本和以任何方式传输文档的功能。XSL 划分为两部分:一种是将 XML 树形结构转化为任意树形结构的转换语言,另一种是描述 XML 树形结构中元素外观的格式化语言。目前支持转换语言的工具比支持格式化语言的工具要多。

第4部分: 补充技术

第4部分介绍了一些基于 XML 的语言和在基本的 XML 顶层之上的语法。XLink 提供多方向的超文本链接,其功能远比简单的 HTML<A>标记强大得多。XPointer 引入了一种新的语法,不仅可以把 URL 的尾部链接到特定文档上,还可以链接到特定文档的个别部分。RDF 是一种用于在 XML 和 HTML 文档中嵌入元数据的 XML 应用程序。元数据是关于文档的信息,例如作者、日期和工作标题(而不是工作本身)等内容。所有这些技术都可以添加到读者自己的基于 XML 的标记语言中,从而扩展其功能和实用性。

第5部分: XML 应用程序

第5部分向读者展示了 XML 在不同领域的应用。XHTML 作为有效的 XML,是对 HTML 4.0 的重组。WML 是一个类似 HTML 的语言,可以使 Web 内容服务于移动电话、PDA、寻呼机、其他存储器、显示器和有限带宽设备。模式是一种基于 XML 的语法,用于描述 XML 文档所允许的内容,大家公认模式比 DTD 的功能更强、扩展性更好。可伸缩的矢量图形(SVG)是 W3C 推荐用于绘图的一种标准 XML 格式。矢量标记语言(VML)是微软专有的 XML 应用程序,Office 2000 和 Internet Explorer 5.0 的矢量图形就使用了该程序。微软的频道定义格式(CDF)是一种基于 XML 的标记语言,可以定义频道,从而将 Web 站点的内容推送到订阅者。最后,为家谱数据研究开发了一个全新的应用程序,这个例子不仅说明如何使用 XML 标记,而且说明选择这些标记的原因和时机。读者通过阅读这些应用程序,将会对在现实世界中如何设计、建立和使用 XML 应用程序有一个很好的感性认识。