

通向全球电子商务的特快车

网页新视界 XML

入门与应用

◎ 陈长念 陈勤意 编著

- XML 与 HTML 有什么区别
- 循序渐进地介绍 XML，让读者系统深入了解 XML 的功能
- 详尽解说 XML 的语法与规范
- XML 相关的技术：CCS, XLL, Namespace, 与 DOM
- 实例示范 XML 在 Internet 上的广泛运用

随书附赠光盘中含有本书中所有精彩实例：W3C 相关 XML 的规范中文
XML 问答：相关 XML 网站：XML 软件：MSXML parser



中国青年出版社



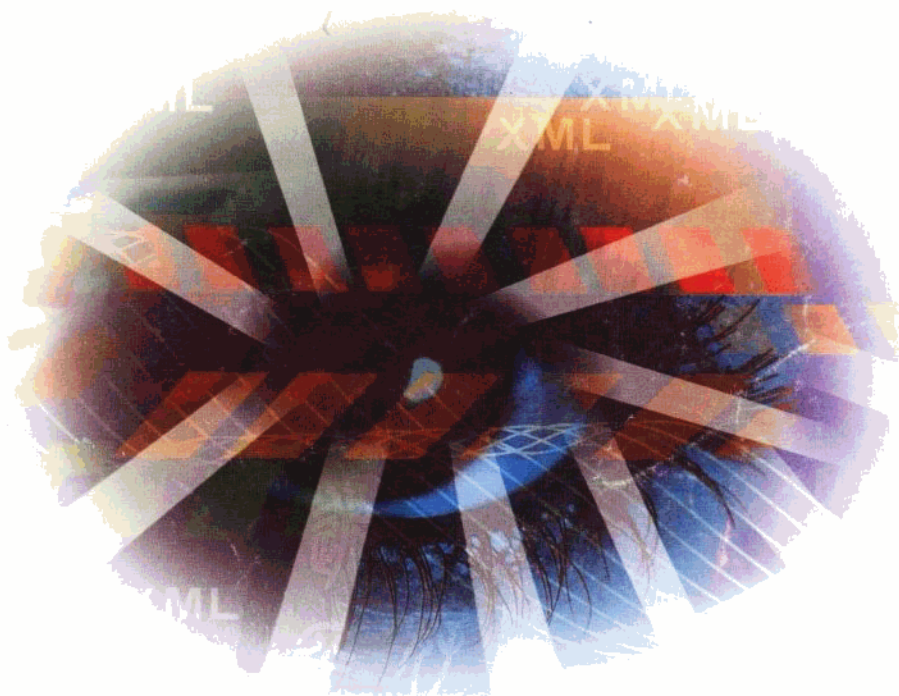
松岗广脑图书资料股份有限公司

网页设计系列

网页新视界 XML

入门与应用

◎ 陈长念 陈勤意 编著



中国青年出版社
CHINA YOUTH PRESS

(京)新登字 083 号

本书繁体字版原名为《网页新视界 XML 入门与应用》，由松岗电脑图书资料股份有限公司出版，版权归松岗电脑图书资料股份有限公司所有。本书简体字中文版由松岗电脑图书资料股份有限公司授权中国青年出版社独家出版。本书封底贴有防伪标签，无标签者即为盗版出版物，不得销售。

版权贸易合同登记号：01-2000-0742

图书在版编目 (CIP) 数据

网页新视界 XML 入门与应用 / 陈长念, 陈勤意编著. —北京: 中国青年出版社, 2000.4
ISBN 7-5006-3774-8

I. 网... II. ①陈...②陈... III. 计算机网络—可扩充语言, XML IV. TP393

中国版本图书馆 CIP 数据核字 (2000) 第 06420 号

样本

TP393.592
7408

策 划: 胡守文
王修文
郭 光
责任编辑: 江 颖
责任校对: 肖新民

书 名:《网页新世界 XML 入门与应用》
编 著: 陈长念 陈勤意
出版发行: 中国青年出版社
地址: 北京市东四十二条 21 号 邮政编码: 100708
电话: (010) 64069368 传真: (010) 64053266

印 刷: 高唐印刷有限公司
开 本: 16 开
版 次: 2000 年 4 月北京第 1 版
印 次: 2000 年 4 月高唐第 1 次印刷
印 数: 1-5000
印 张: 24.5
定 价: 48.00 元(随书附赠 CD)

序

继 HTML 与 JAVA 在 Internet 上刮起旋风之后, XML 是目前在 Internet 上最热门的话题, 几乎每天在 Internet 上都有 XML 最新的应用被提出, 而各大软件厂家也纷纷提出支持 XML 的应用, 且在 Internet 上专门讨论 XML 技术的网站也日益增多, 一些相关 XML 的免费使用软件也越来越多, 当然介绍 XML 的书籍也相当的多, 各种迹象都显示出 XML 是未来 Internet 上信息类型的主流, 所以不管你在 Internet 的应用中是扮演哪种角色, 是服务的提供者、网页设计者、网站维护者还是一般的使用者, 你都需要认识 XML, 而本书的内容是针对 XML 作一深入介绍, 并实际示范目前 XML 在 Internet 上的应用, 让读者能了解 XML 在 Internet 所扮演的角色并懂得其应用。

XML 的全文是 eXtensible Markup Language, 中文可译为可扩展的标记语言, 全球信息网协会(W3C)已于 1998 年 2 月通过 XML 1.0 规范的审核, 且一些与 XML 相关的规范如 XSL、XLink、XPointer、XHTML... 等等, 都已在进行审核之中, 而本书除了介绍 XML 外, 对于上述与 XML 相关的规范也会一并作介绍, 让读者可以完全掌握 XML 未来的发展与应用。

虽然 XML 目前在 Internet 上的应用并不比 HTML 来得广泛, 但笔者认为这是时间上的问题, 就好比 DOS 的操作系统最终还是要迈入 WINDOWS 操作系统的, 这是一种里程碑, 但这并不表示 DOS 操作系统将完全被遗弃, 目前 WINDOWS 操作系统仍保留有 DOS 的操作环境; 相同地, 这也不代表 XML 可以完全取代 HTML, 而是 XML 可以补足 HTML 的缺憾, 能完成 HTML 无法或不易完成的事。

所以不认识 HTML 的人可以直接学习 XML，而已学过 HTML 的人更需要认识 XML，目前国内有关 HTML 的中文书籍已是相当的多，但可惜的是直至笔者撰写本书之前，国内尚无一本专门介绍 XML 的中文书籍，这也是笔者撰写本书最大的原动力，希望能将笔者使用 XML 的经验与读者分享，并催化加速 XML 能在国内被广泛熟知，当然笔者才疏学浅，也希望各位读者先进能不吝赐教。

最后仅以本书献给老陈刚刚出生的女儿香湘，祝她健康、平安、快乐！

陈长念 陈勤意

2000 年 3 月

目 录

PART 0 前言

0-1 本书摘要.....	2
0-2 本书光盘使用说明.....	4

PART I XML 基础篇

Chapter 1 标记语言

1-1 认识标记语言.....	18
1-1-1 何谓标记语言	18
1-1-2 RTF 标记语言.....	20
1-2 标准通用标记语言.....	27
1-2-1 SGML 功能与用途.....	27
1-2-2 SGML 优点.....	29
1-2-3 SGML 缺点.....	31
1-3 超链接标记语言.....	32
1-3-1 HTML 功能与用途.....	32
1-3-2 HTML 优点.....	38
1-3-3 HTML 缺点.....	40
1-4 可扩展标记语言.....	42
1-4-1 XML 功能与用途.....	42
1-4-2 XML 与 SGML	45
1-4-3 XML 与 HTML	46
本章习作.....	49

Chapter 2 认识 XML

2-1	制定 XML 的目标	52
2-2	撰写 XML 文件的工具	56
2-2-1	介绍如何使用记事本来编辑 XML 文件	57
2-2-2	介绍有哪些专门用来编辑 XML 文件的软件	59
2-3	XML 文件的解析器	62
2-3-1	解析器的功能作用	62
2-3-2	解析器种类	63
2-3-3	用 IE5.0 来解析 XML 文件	64
2-3-4	MSXML 安装与使用	68
2-4	XML 应用架构说明	73
2-4-1	XML 文件与样式表	74
2-4-2	Intranet 与 Extranet 应用	75
2-4-3	不同数据库间的信息交换	76
2-5	W3C 所制定与 XML 相关的标准	74
2-5-1	推荐使用的标准	78
2-5-2	准备推荐使用的标准	79
2-5-3	尚在草案制定阶段的标准	80
2-6	XML 邮递论坛	81
	本章习作	84

Chapter 3 XML 文件的结构

3-1	XML 文件的逻辑结构	86
3-1-1	文件档头	86
3-1-2	文件本体	89
3-2	XML 文件的实体结构	91
3-3	XML 基本语法介绍	92
3-3-1	控制标记与命名	93
3-3-2	开始标记与结束标记	95
3-3-3	属性设定	96
3-3-4	空的控制标记	97
3-3-5	备注标记	98
3-3-6	文字 CDATA 标记	100

3-4 XML 文件可以使用中文	103
3-4-1 文件内容可以包含中文	103
3-4-2 中文的控制标记	105
本章习作	108

Chapter 4 Well-Formed 的 XML 文件

4-1 XML 文件的宣告	110
4-2 XML 文件中只能有一个根节点	112
4-3 对称的控制标记	114
4-4 空的控制标记	115
4-5 满足巢状的排列	116
4-6 英文大小写有差异	118
4-7 属性值的设定	120
4-8 特殊字符	121
本章习作	124

Chapter 5 文件格式定义

5-1 为何要使用文件格式定义	126
5-1-1 何谓文件格式定义	126
5-1-2 使用文件格式定义的好处	129
5-1-3 Validating XML 文件	130
5-2 如何调用文件格式定义	131
5-2-1 内部文件格式定义	131
5-2-2 外部文件格式定义	134
5-3 元素的设定与使用	137
5-3-1 元素	137
5-3-2 空元素的设定	142
5-3-3 不限控制标记出现的次序与次数	144
5-3-4 根元素、父元素与子元素	146
5-4 元素数目设定	148
5-4-1 宣告元素可以出现的次数	148
5-4-2 选择性元素	150
5-4-3 混合型元素	152
5-4-4 可以有递归的结构	154

5-5	实体的宣告与使用	155
5-5-1	实体的分类	155
5-5-2	内部实体的设定与使用	157
5-5-3	外部实体的设定与使用	160
5-5-4	实体的应用范围	163
5-5-5	实体的参数型调用	166
5-6	IGNORE 与 INCLUDE	168
5-7	文件格式定义应用	171
5-7-1	内部 DTD 结合外部 DTD	171
5-7-2	整合数个外部 DTD	173
	本章习作	177

Chapter 6 元素的属性

6-1	元素的属性介绍	180
6-1-1	认识属性	180
6-1-2	在 DTD 中如何设定元素的属性	181
6-2	属性的内定值	184
6-2-1	需要被设定的属性	184
6-2-2	视需求来设定的属性	186
6-2-3	固定值的属性	187
6-2-4	有内定值的属性	189
6-3	属性类型	191
6-3-1	CDATA	192
6-3-2	Enumerated	192
6-3-3	NMTOKEN	194
6-3-4	NMTOKENS	196
6-3-5	ID	198
6-3-6	IDREF	199
6-3-7	IDREFS	201
6-3-8	ENTITY	202
6-3-9	ENTITIES	203

6-3-10 NOTATION	205
6-4 内定的属性.....	206
6-4-1 空白格的处理	207
6-4-2 设定文字的语言	209
本章习作.....	211

PART II XML 应用篇

Chapter 7 IE 浏览器的推播频道

7-1 IE 浏览器推播频道介绍	216
7-1-1 何谓推播频道	216
7-1-2 建立推播频道的步骤	218
7-2 推播频道的 DTD 介绍.....	222
7-2-1 主要元素介绍	222
7-2-2 设定时程表元素	226
7-2-3 其它元素	232
7-3 IE 浏览器推播频道制作	236
7-3-1 决定推播频道的内容	236
7-3-2 制作推播频道	238
7-3-3 上载频道相关数据	240
7-3-4 下载推播频道	242
本章习作.....	244

Chapter 8 CSS 样式表

8-1 何谓 CSS 样式表.....	246
8-1-1 介绍 CSS 样式表	246
8-1-2 CSS 样式表的好处与限制.....	249
8-2 如何使用 CSS 样式表.....	251
8-2-1 在 XML 文件中如何调用 CSS 样式表	251
8-2-2 空的 CSS 样式表	253

8-2-3	配合 DTD 来调用 CSS 样式表	255
8-2-4	CSS 属性介绍	256
8-3	CSS 样式表应用	259
8-3-1	DISPLAY 属性的设定	260
8-3-2	多层元素的属性设定	262
8-3-3	版面设定	265
8-3-4	设定外框	267
8-3-5	使用不同的样式表	271
8-3-6	多笔信息的显示	273
	本章习作	277

Chapter 9 XSL 样式表

9-1	何谓 XSL 样式表	280
9-2	如何使用 XSL 样式表	281
9-2-1	在 XML 文件中如何调用 XSL 样式表	281
9-2-2	IE5.0 与 XSL	282
9-2-3	XSL 元素介绍	284
9-3	XSL 基本元素	286
9-3-1	基本的样式表	286
9-3-2	设定排版属性	293
9-3-3	xsl:comment	295
9-3-4	xsl:for-each	298
9-3-5	HTML 的表格控制标记	302
9-3-6	显示额外的信息	304
9-3-7	xsl:apply-templates	307
9-3-8	xsl:copy	311
9-4	XSL 中条件式元素	313
9-4-1	条件设定	313
9-4-2	多重条件设定	319
9-5	建立控制标记与属性	327
9-5-1	建立属性	327
9-5-2	建立控制标记	333
	本章习作	337

Chapter 10 XML 进阶应用

10-1 Namespaces.....	340
10-1-1 为何需要使用 Namespaces.....	340
10-1-2 如何宣告.....	341
10-1-3 XSL 与 Namespaces 应用.....	344
10-1-4 如何在 XML 中引用 HTML.....	345
10-2 DOM.....	349
10-3 XLink.....	353
10-3-1 XLink 介绍.....	354
10-3-2 Simple Link.....	356
10-3-3 Extended Link.....	360
10-4 XPointer.....	362
10-4-1 XPointer 介绍.....	363
10-4-2 XPointer 寻址方法.....	364
本章习作.....	369

Appendix A

PART 0

前言

摘要

本书将分两部分来为读者深入介绍 XML, Part I 是 XML 基础篇, 包含本书的第一章到第六章; 而 Part II 是 XML 应用篇, 包含本书第七章到第十章, 全书共计十章。

0-1 本书摘要

本书将分两部分来为读者深入介绍 XML，Part I 是 XML 基础篇，包含本书的第一章到第六章；而 Part II 是 XML 应用篇，包含本书第七章到第十章，全书共计十章，有关 Part I 与 Part II 摘要如下：

Part I XML 基础篇

以循序渐进的方式来介绍 XML，让读者深入了解何谓 XML 及其功用，对于 XML 的语法与规范有详尽的解说，完成本篇后读者将有建立自己的标记语言与 XML 文件的能力。

第一章 标记语言

介绍何谓标记语言及其功用，其中包括介绍 SGML 与 HTML 的应用情况、优缺点，并说明 XML 为何能成为新一代的标记语言。

第二章 认识 XML

介绍 XML 的发展过程与现况，包括制定 XML 的目标是什么、发展与使用 XML 的软件工具介绍、目前 XML 被应用的情况及目前在 Internet 上有哪些专门讨论 XML 的邮件论坛。

第三章 XML 文件的结构

本章将深入介绍 XML 文件的结构，XML 文件在结构上可以用逻辑与实体结构来说明，其实两者对 XML 文件来说是一体两面，是同时存在的，如果将 XML 文件比拟作一栋建筑物，逻辑结构就像是建筑物的蓝图，实体

结构就像是该建筑物的实质构造，是实际包括使用那些材料所组成的。

第四章 Well-Formed 的 XML 文件

构成一份 XML 文件必须要有一些规格限制，而最基本的限制就是一份 XML 文件必须是 well-formed，在本章中共有 8 个小节，在每个小节中将以正反范例来介绍一种满足 well-formed 格式的条件，若一份 XML 文件能同时满足这 8 个条件，则该份 XML 文件可以称之为 well-formed。

第五章 文件格式定义

本章将深入介绍如何撰写文件格式定义(Document Type Definition, DTD)，并说明 DTD 的种类、DTD 如何与 XML 文件结合使用。另外在本章也将让读者了解何谓 Validating 的 XML 文件，DTD 在 XML 文件中的使用是何有重要性。

第六章 元素的属性

本章将继续介绍如何撰写文件格式定义(DTD)，但重点是着重在 DTD 文件中元素的属性介绍，包括如何在元素中宣告属性、属性类型与属性内定值的种类。

Part II XML 应用篇

在完成 XML 基础篇的介绍后，读者一定很好奇 XML 可以做怎样的应用，因为单纯的一个 XML 文件只是包含一些特定格式的信息，必须配合特殊软件来解读应用，才能真正让读者发现 XML 文件其广泛的应用范围与弹

性。

在本篇中将介绍 XML 在推播频道与网页设计上的应用，其中推播频道已正式在 IE4.0 以上的浏览器中应用，在第七章中也将详细介绍推播频道如何使用 XML 规范来制定一个文件格式定义。另外有关网页设计方面，在第八、九章中将分别介绍 CSS 与 XSL 如何与 XML 配合来完成网页的显示，在第十章中将介绍一些与 XML 相关的重要应用规范，这些规范与网页设计也是息息相关的。

第七章 IE 浏览器的推播频道

Channel Definition Format 简称 CDF 由 Microsoft 和 Marimba 所开发出来的一种标准，CDF 即是一种 XML-Based 的标记语言用来定义推播频道，主要应用在线上出版和 Push 技术上。目前 Microsoft 的 IE 4.0 版以上浏览器和 PointCast 皆采用 CDF，本章将深入介绍推播频道的功能与制作。

第八章 CSS 样式表

本章将深入介绍 CSS 样式表如何被 XML 文件套用，并将 XML 文件的内容显示在浏览器中。

第九章 XSL 样式表

XSL 本身即是一种 XML 标记语言的应用，本章将深入介绍如何结合 XML 文件与 XSL 样式表，将 XML 文件内容通过 XSL 的排版设定显示在浏览器中。

第十章 XML 进阶应用

本章将介绍其它一些与 XML 在网页设计上相关的技术,包括 Namespaces、DOM、XLink 与 XPointer。

0-2 本书光盘使用说明

本书光盘中的信息可分为网页类与应用软件类,所谓的网页类就是该类信息可以使用浏览器来直接浏览阅读,而应用软件类则是一些与使用 XML 相关的软件。

网页类

用鼠标点选桌面“我的电脑”后,再点选本书的光盘片后,如图 0-1 所示,共有 4 个目录与 index.htm 文件,其中 index.htm 即是本书的首页。

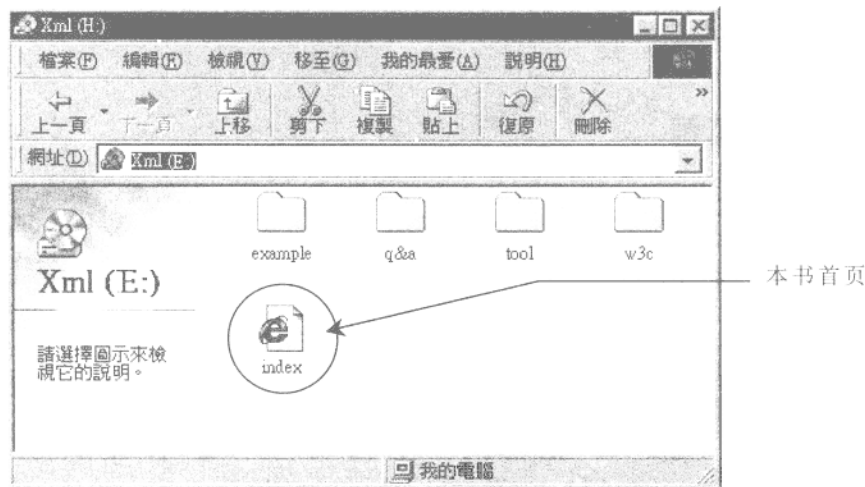


图 0-1