高等院校计算机应用系列教材

XML 基础教程

(第二版)(微课版)

高宇飞　主编

清華大學出版社

北　京

内 容 简 介

本书从初学者角度出发，以通俗易懂的语言，详尽丰富的实例，介绍了 XML 相关的各种主要技术。书中不仅详细阐述了 XML 的基本概念、语法规则、文档类型定义、层叠样式表、可扩展样式表、解析器和数据库的集成等知识，还通过一个综合案例演示了 XML 在实际项目开发中的应用。

本书注重基础、讲究实用、力求由浅入深，在讲解基本概念和基础知识的同时给出了大量实例，便于读者掌握所学的内容。每章还包括小结和习题，便于读者巩固所学的知识。本书可作为高等院校软件工程、计算机科学与技术等相关专业的研究生参考用书，也可作为相关专业的高年级本科教材，还可作为初学者学习 XML、Android 移动应用开发、Java EE 开发的培训教材。

本书配套的电子课件、实例源文件、习题答案可以到 http://www.tupwk.com.cn/downpage 网站下载，也可以扫描前言中的二维码获取。读者扫码前言中的视频二维码可以直接观看教学视频。

图书在版编目(CIP)数据

XML 基础教程：微课版 / 高宇飞主编. —2 版. —北京：清华大学出版社，2022.7
高等院校计算机应用系列教材
ISBN 978-7-302-61095-3

Ⅰ. ①X… Ⅱ. ①高… Ⅲ. ①可扩充语言—程序设计—高等学校—教材 Ⅳ. ①TP312

中国版本图书馆 CIP 数据核字(2022)第 101037 号

责任编辑：胡辰浩
封面设计：高娟妮
版式设计：孔祥峰
责任校对：成凤进
责任印制：朱雨萌

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　邮　　编：100084
社 总 机：010-83470000　　邮　　购：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：三河市科茂嘉荣印务有限公司
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张**：15.25　　**字　　数**：390 千字
版　　次：2015 年 4 月第 1 版　　2022 年 8 月第 2 版　　**印　　次**：2022 年 8 月第 1 次印刷
定　　价：69.00 元

产品编号：090788-01

前　言

在以计算机与互联网技术为代表的IT时代，各种各样的新技术如雨后春笋般涌现，然而真正能够历经磨炼生存下来的却寥寥无几。毫无疑问，XML便是其中的佼佼者。XML是SGML的一个子集，它保留了SGML的灵活性，去掉了其复杂性。XML诞生不久，很快便获得了巨大的成功，XML标准开始突飞猛进地发展，大批的软件开发商争先恐后地采纳这个标准，这一切令人赞叹不已。如今，XML在IT领域已经拥有不可动摇的地位，一些重要的应用程序都使用XML来保存它们的配置文件或数据文件。

XML是由W3C定义的一种语言，是表示结构化数据的行业标准。XML在电子商务、移动应用开发、Web Service、云计算等技术和领域中起着非常重要的作用。一些名人曾这样评论XML。

- 微软总裁比尔•盖茨：XML将为每一种流行的编程语言带来一场语言革命，其影响力甚至超过HTML为世界带来的影响。
- 微软CEO史蒂夫•鲍尔默：XML的出现，对于信息技术的影响不亚于GUI和浏览器。
- IBM资深专家Goldfarb：我为XML感到骄傲，WWW正在转为以XML为基础。

XML是未来的发展趋势，无论是网页设计师还是网络程序员，都应该及时学习和了解，一味等待只会让你失去机会。

应该学习和掌握XML的理由如下。

- XML是一门较新的技术。
- XML是最前沿的技术。
- XML是应用广泛的技术，其发展前景无可限量。
- XML是一门综合性很强的技术。

XML越来越受追捧，关于XML的基础教程也随处可见，可是一大堆的概念和术语往往让人望而生畏。有些图书起点太高，初学者难以理解基本概念，一开始学习就困难重重，容易产生厌倦心理而放弃；有的图书又过于简单，读者学完之后还是不会做实际项目，不能达到提升自己技能的目的。

概括起来，本书具有以下主要特点。

- 注重基础，讲究实用，力求从入门到精通。
- 充分体现案例教学。本书以易学易用为重点，例子实用、知识丰富、步骤详细、学习效率高，特别适合入门者。
- 配有电子课件、教学视频、习题答案和实例源文件。本书的所有示例均在XML Spy 2013开发环境下调试通过，读者可直接下载所有例子的源程序，并通过教材中介绍的步骤学习要点。

本书在讲述 XML 基本概念的基础上，系统地介绍了 XML 技术中已成熟的标准和应用技术，并给出了基于 XML 的应用实例。全书共分为 10 章，各章的主要内容如下。

第 1 章是 XML 简介，讲述标记语言的发展、HTML 的局限性、XML 的实现机制、XML 的优势与特点，并给出了 XML 文档范例。这一章还用不少的篇幅介绍了 XML 技术的应用领域与应用前景，以及与 XML 相关的各种技术。

第 2 章讲解 XML 的语法，包括 XML 文档的构成、XML 文档的声明与注释、XML 元素的组成与命名、XML 元素属性的定义规则、特殊的 CDATA 文本段、XML 命名空间的概念与应用等。XML 的语法并不复杂，但在编写 XML 文档时必须遵循这些语法规则，只有这样才能编写出格式良好的 XML 文档。

第 3 章讲解文档类型定义 DTD，介绍了 DTD 的基本结构，重点阐述如何使用 DTD 为 XML 文档建立语义约束，包括如何在 DTD 中定义元素及元素类型，分析 DTD 所支持的各种属性类型，说明如何在 DTD 中定义各种实体，指出 DTD 的局限性及现状。

第 4 章讲解描述和约束 XML 文档的语言——XML Schema。对比 DTD 中存在的缺陷引出了 Schema，以一个 Schema 文档为例，介绍 Schema 的基本结构，详细分析 Schema 中的简单类型和复杂类型，说明如何进行数据类型的定义、元素的定义和属性的定义，分析 Schema 命名空间的作用，介绍验证 XML 文档有效性的两种方法。

第 5 章介绍如何使用 CSS(层叠样式表)来格式化输出 XML 文档的内容。XML 文档本身只包含数据而不包含这些数据的显示格式信息，然而利用简单的 CSS 技术就能实现将 XML 文档中的数据以设计者所设定的各种格式在浏览器中显示出来。

第 6 章讲解 XSL(可扩展样式表)技术，利用该技术不仅能够把 XML 文档转换为 HTML 文档，实现在浏览器中的格式化显示，还可以将 XML 文档转换为其他各种基于文本的文档，以实现跨平台的数据共享和交换。

第 7 章详细展示 XML 文档的解析过程，包括 DOM 树模型、DOM 的结构、DOM 基本接口、DOM 的节点访问和 DOM 对 XML 文档的相关操作等内容。DOM 解析器的主要功能是检查 XML 文件是否有结构上的错误，剥离 XML 文件中的标记，读出正确的内容，并交给下一步应用程序处理。

第 8 章介绍一种高效的解析器——SAX，包括 SAX 的优缺点、工作机制、事件处理器、SAX 事件、常用接口、回调方法、SAX 错误信息和 SAX 对 XML 文档的相关操作。在这一章中还比较了 SAX 与 DOM 两种截然不同的解析方式，并给出了将两者结合应用的具体实例。

第 9 章介绍 XML 与关系数据及关系数据库的集成，阐述数据库技术的发展、XML 的数据交换及存取机制、在数据库技术中引入 XML 的原因以及二者的结合对数据交换的影响，并全面介绍.NET 平台下 XML 与关系数据库系统互换数据所采用的各种技术，以及 SQL Server 2019 对 XML 的支持。

第 10 章通过一个综合性的实例，系统介绍 DOM、SAX、CSS 等多种 XML 技术的应用，演示在.NET 平台下利用 XML 进行实际项目开发的完整过程。

本书从 XML 的基础知识讲起，语言通俗易懂，并配有丰富的实例和插图，使读者对每一章所讲述的内容都能有深刻的理解，十分适合初学者和有一定 XML 基础的人员使用。

本书由高宇飞主编，参与本书编写的人员还有杨亚锋、刘皓雯、徐静、谢素祯、王震源、张吉涛、彭少康、祁子豪、陈震、李天、马自行、宋嘉强、王玉森、王兆楠、薛红秋、王天宝、李世博和王向杰等。同时，对清华大学出版社表示感谢。

由于作者水平有限，书中难免有不足之处，恳请专家和广大读者批评指正。在本书的编写过程中参考了一些相关文献，在此向这些文献的作者深表感谢。我们的电话是 010-62796045，邮箱是 992116@qq.com。

本书配套的电子课件、实例源文件、习题答案可以到 http://www.tupwk.com.cn/downpage 网站下载，也可以扫描下方的二维码获取。扫码下方的视频二维码可以直接观看教学视频。

配套资源　　扫一扫

扫描下载　　看视频

编者

2022 年 3 月

目　录

∞ 第 1 章 ∞

XML简介

在互联网的发展历史上，有两种非常核心的技术，分别是 Java 和 XML。Java 提供了程序代码的平台无关性；而 XML 则保证了数据的平台无关性，被誉为因特网上的世界语，已成为 Web 应用中数据表示和数据交换的标准。但是，人们对 XML 的认识远远没有对 HTML 的认识彻底和清晰。那么，究竟什么是 XML？XML 和 HTML 有什么区别？它们的本质区别是什么？另外，由于 XML 的优越性及 XML 的不断发展壮大，XML 的标准和规范不断变化，了解这些标准的来龙去脉以及它们之间的关系，对于掌握 XML 至关重要。

本章首先介绍标记语言的发展历史，在与有关标记语言比较的基础上，引出 XML，然后对 XML 的特点、作用，以及与之相关的技术进行简要介绍。通过本章的学习，读者将了解到 XML 技术的具体含义及其广阔的应用前景。

本章的学习目标：

- 掌握 XML 的特点
- 理解 XML 与 HTML 的区别
- 了解 XML 的应用领域
- 掌握 XML 的技术规范
- 熟悉 XML 文档的编辑软件

1.1 XML 的产生

XML 的全称是 Extensible Markup Language，即可扩展标记语言，它是 SGML 的一个子集，现在广为使用的 HTML 也是 SGML 家族中的一员。

HTML、XML 以及 SGML 都属于标记语言。标记语言不同于 Java、C 这样的编程语言，它本身并无任何“动作行为”，比编程语言简单得多。标记语言只是用一系列约定好的标记来对电子文档进行标记，从而为电子文档额外增加语义、结构和格式等方面的信息。

1.1.1 SGML 的诞生

在 20 世纪 60 年代，IBM 的研究人员提出在各文档之间共享一些相似的属性，如字体大小和版面。IBM 设计了一种文档系统，通过在文档中添加标记，来标识文档中的各种元素，IBM 把这种标记语言称作通用标记语言(Generalized Markup Language)，即 GML。

在当时的信息交换过程中，经常会遇到数据格式不同的问题，而且随着网络技术的不断发展，这一问题日益严重，制约了人们的信息交流。经过若干年的发展，GML 在 1986 年演变成一个国际标准(ISO 8879)，并被称为 SGML(Standard Generalized Markup Language)，即标准通用标记语言。SGML 是一种定义电子文档结构和描述其内容的国际标准语言，是所有电子文档标记语言的起源，早在 Web 出现之前就已存在。SGML 具有良好的扩展性和可移植性，在任何一种环境下都可以正常使用。但 SGML 强大功能的背后是它的复杂度太高，不适合网络的日常应用。另外，SGML 价格昂贵，开发成本高。更为重要的是，它不被主流浏览器厂商所支持。这些原因均使得 SGML 的推广受到了阻碍。标记语言的发展历史如图 1-1 所示。

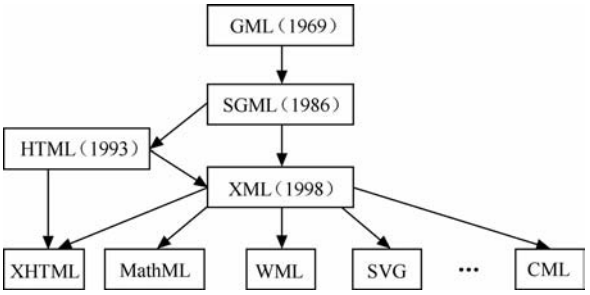

图 1-1　标记语言的发展历史

1.1.2　什么是 XML

超文本标记语言(Hypertext Markup Language，HTML)是目前网络上应用最广泛的语言，也是构成网页文档的主要语言。HTML 中的标记都是在 HTML 5 中规范和定义的，而 XML 允许用户自己创建这样的标记，所以说 XML 具有可扩展性。XML 文件是由标记以及它所包含的内容构成的文本文件，这些标记可自由定义，其目的是使得 XML 文件能够很好地体现数据的结构和含义。W3C 推出 XML 的主要目的是使 Internet 上的数据交互更方便，让文件的内容更易懂。

XML 同 HTML 一样，都源自 SGML。SGML 十分庞大，既不容易学习，又不容易使用，在计算机上实现也十分困难。鉴于这些因素，Web 的发明者(欧洲粒子物理研究中心的研究人员)根据当时(1989 年)的计算机技术，开发了 HTML。

HTML 只使用了 SGML 中很少的一部分标记，如 HTML 4.0 中只定义了 70 余种标记。为了便于在计算机上实现，HTML 规定的标记是固定的，即 HTML 语法是不可扩展的。HTML 这种固定的语法使它易学易用，在计算机上开发 HTML 的浏览器也十分容易。正是由于 HTML 的简单性，使得基于 HTML 的 Web 应用得到了极大的发展。

但随着 Web 应用的不断发展，HTML 的局限性也越来越明显。首先，HTML 可以指定一个文档的内容和格式，但不能指定文档的结构。也就是说，HTML 是面向表示而非面向结构的标记语言，只能用来告诉浏览器如何在网站上显示信息。其次，HTML 只能应用于信息的显示，它可以使文本加粗，以斜体或下画线形式显示，但它几乎没有语义结构。HTML 对数据的显示是按照布局而非按照语义。随着网络应用的飞速发展，各行各业对各种信息有着不同的需求，这些不同类型的信息未必都是以网页的形式显示出来。例如，当通过搜索引擎进行数据搜索时，按照语义而不是按照布局来显示数据显然更具优势。另外，HTML 的可扩展性较差，HTML 中

标记的名称是固定不变的，因而其提供的功能与使用的属性也是固定的。所以 HTML 不允许网页设计者自行创建标记。例如，HTML 文档包括了格式化、结构和语义的标记。<B>就是 HTML 中的一种格式化标记，它使其中的内容变为粗体；<TR>也是 HTML 中的一种结构标记，指明内容是表格中的一行。也就是说，HTML 不是一种元语言，不能创建某一特定领域的标记集。虽然作为一般的应用，HTML 已经够用了，但科学家无法用 HTML 书写数学公式、化学方程式以及分子晶体结构，这样使它的发展受到了极大的限制。

总而言之，HTML 的缺点使其交互性差，语义模糊。随着互联网应用的发展，HTML 越来越难以满足网络数据交互和业务集成的需求。

有人建议直接使用 SGML 作为 Web 语言，这固然能解决 HTML 遇到的困难。但是 SGML 过于庞大，用户学习、使用不方便尚且不说，仅是熟练使用 SGML 的浏览器就非常困难。于是自然想到了使用 SGML 的子集，这样既方便使用又容易实现。正是在这种形势下，Web 标准化组织 W3C 建议使用一种精简的 SGML 版本——XML 由此应运而生。

XML 是 SGML 的一个精简子集，其复杂度大约只有 SGML 的 20%，但却是有 SGML 80% 的功能，因此它一经推出即受到用户的欢迎。XML 保留了 SGML 的可扩展功能，这使 XML 从根本上有别于 HTML。XML 是一种元标记语言，要比 HTML 强大得多，它的标记不再固定，而是需要用户根据描述数据的需求自己定义。这些标记必须根据某些通用的规则来创建，但是标记的意义具有较大的灵活性。

例如，在 HTML 中，一首歌可能是用定义标题标记<dt>、定义数据标记<dd>、无序列表标记<ul>和列表项标记<li>来描述的。但是事实上这些标记没有一个是与音乐有关的。用 HTML 定义的歌曲如下。

```
<dt>金曲 TOP1
<dd>春暖花开
<ul>
    <li>词：梁芒
    <li>曲：洪兵
</ul>
```

而在 XML 中，同样的数据可能标记如下。

```
<song>金曲 TOP1
<title>春暖花开</title>
<composer>洪兵</ composer >
<lyricist>梁芒</lyricist>
</song>
```

这段代码中没有使用通用的标记如<dt>、<li>等，而是使用了更有意义的标记，如<song>、<title>、<composer>等。这种用法使源代码易于阅读，使人能够读懂代码的含义。

XML 具有以下特点。

- XML 描述的是结构和语义，而不是格式化。
- XML 将数据内容和显示格式相分离。
- XML 是元标记语言。XML 的标记不是预先定义好的，而是自定义的。
- XML 是自描述语言。XML 使用 DTD 或者 Schema 后就是自描述的语言。XML 文档通

常包含一个文档类型声明，因而它是自描述的。不仅人能读懂 XML 文档，计算机也能处理。

- XML 是独立于平台的。
- XML 不进行任何操作。
- XML 具有良好的保值性。XML 良好的保值性和自描述性使它成为保存历史档案(如政府文件、公文、科学研究报告等)的最佳选择。

XML 标准的发展没有 HTML 那样迅速，直到 1998 年 2 月，W3C 才发布了 XML 1.0 推荐标准，在 2000 年 10 月发布 XML 1.0 推荐标准的第二版，在 2004 年 2 月，发布了 XML 1.0 推荐标准的第三版和 XML 1.1 的推荐标准。目前最新的 XML 版本是 2006 年 8 月发布的 XML 1.1 的推荐标准，不过目前大多数的应用程序遵循的还是 W3C 于 2000 年 10 月 6 日发布的 XML 1.0 标准。

1.1.3 XML 和 HTML 的区别

从前面的介绍中，我们可以感觉到 HTML 和 XML 的明显区别。HTML 标记用途很简单，也很明确，就是使用 HTML 标记创建的文档可以用浏览器显示相似的内容，并显示美观的网页编排。而 XML 则属于一种文档格式的革命，它能让用户自定义文档结构，给予文档一种全新的生命，让计算机能够读懂文档。XML 的设计目的是在不同的计算机平台和不同的计算机程序间方便、平稳地交换数据，从而提高处理数据的效率和灵活性。

下面对二者之间的差异进行比较。

(1) XML和HTML都源自SGML，它们都含有标记，有着相似的语法，区别在于：HTML不具有扩展性，它用固有的标记来描述、显示网页内容。例如，＜H1>是第一级标题标记，有固定的尺寸——20 磅的Helvetica字体的粗体。如果HTML没有定义用户所需的标记，用户就束手无策了，只能等待HTML的下一个版本，希望在新版本中能包括所需的标记。而XML是元标记语言，可用于定义新的标记语言。如果将HTML比作是在织毛衣，那么XML就是关于如何织毛衣的指导书。学会XML，用户不仅可以织毛衣，还可以织袜子、手套等。

(2) HTML 的核心不是为了体现数据的含义，而是为了体现数据的显示格式。HTML 网页将数据和显示混在一起，而 XML 则将数据和显示分隔开。XML 的核心是描述数据的组织结构，让 XML 可以作为数据交换的标准格式。由于 XML 文档本身不受表现形式的束缚，只要对 XML 文档进行适当的转换，就可以将其变成不同的形式，如网页、PDF 文档和 Word 文档等，可以达到“一次编写，多处使用”的目的，提高了内容的可重用性。

(3) 吸取 HTML 松散格式带来的经验教训，XML 一开始就要求遵循语法规则，编写的文档要具有“良好的格式”。下面这些语句在 HTML 中随处可见。

```
＜b>＜ i>sample＜ /b>＜ /i>1-
＜td>sample＜ /TD>
＜font color=red>samplar＜ /font>
```

而在 XML 文档中，上述几种语句的语法都是错误的。XML 严格要求嵌套、配对和遵循 DTD 的树结构。

XML 和 HTML 的更多区别在表 1-1 中进行了详细对比。

表 1-1　XML 和 HTML 的区别

比较内容	HTML	XML
是否预置标签	预置大量标签	自定义标签
可扩展性	不具有可扩展性	是元标记语言，可用于定义新的标记语言，具有很好的可扩展性
侧重点	侧重于如何表现信息	侧重于传输和存储数据，核心是数据本身
语法要求	松散、不严格	严格要求嵌套、配对，并遵守 DTD 或 Schema 定义的语义约束
可读性及可维护性	难以阅读和维护	结构清晰，便于阅读和维护
数据和显示的关系	数据与显示混为一体，难以分离	数据与显示分离
与数据库的关系	与数据库没有关系	与关系数据库的数据表对应，可进行转换
是否区分大小写	大部分浏览器不区分大小写	严格区分大小写
编辑工具	文本编辑工具，大量所见即所得的编辑器(如 Dreamweaver)	文本编辑工具，大量 XML 编辑器(如 XML Spy)
处理工具	任何浏览器均可	需要专门的程序进行处理

下面再通过具体的示例将 HTML 和 XML 进行对比，例 1-1 中的 example1-1.html 是一个简单的 HTML 文件。

【例 1-1】example1-1.html 文件的源代码如下。

```
<html>
<head>
    <title>订单信息</title>
</head>
<body>
    <h1>订单号：1001</h1>
    <h2>商品名称：运动服</h2>
    <h2>单价：200 元</h2>
    <h2>数量：15 双</h2>
</body>
</html>
```

上面的标记，如<html>、<head>、<body>、<h1>等都是固定的，而在创建 XML 文档时，则可以由用户自定义各种标记并以任何名称命名它们。与之对应的 XML 文件 example1-1.xml 如下。

```
<?xml version="1.0" encoding="gb2312"?>
<订单>
    <订单号>1001</订单号>
    <商品名称>运动服</商品名称>
```

```
        <单价>200</单价>
        <数量>15</数量>
    </订单>
```

从 example1-1.xml 中可以很清楚地看出数据的组织结构，所以 XML 文档其实什么都不做，它只是用 XML 标记存储信息的文件。

总之，XML 使用一个简单而又灵活的标准格式，为基于 Web 的应用提供了一种描述数据和交换数据的有效手段，但 XML 并非是用来取代 HTML 的。事实上，它们是基于两个不同的目标而开发的。HTML 着重于描述如何将文件显示在浏览器中，XML 和 SGML 相近，着重于描述如何将文件以结构化的方式表示。就网页显示功能来说，HTML 比 XML 要强大；但就文件的应用范畴来说，XML 比 HTML 的应用要广泛。

1.2 XML 的现状与发展

XML 具有许多优良的特性，并且使用方便，因此越来越受青睐。目前，许多大公司和开发人员已经开始使用 XML，包括 B2B 在内的很多优秀应用都已经证实了 XML 将会改变今后创建应用程序的方式。当然，XML 的意义远非如此，其潜在的影响是深远的。

1.2.1 XML 的应用领域

XML 在实际使用的过程中发挥着巨大的作用。目前，越来越多的行业开始使用 XML 来实现特定的功能。

XML 的用途主要包括以下几个方面。

1. 从 HTML 中分离数据

在不使用 XML 时，数据必须存储在 HTML 文件内；使用 XML 后，数据就可以存放在分离的 XML 文档中。HTML 只需要实现数据的显示和布局，这样当数据发生变动时不会导致 HTML 文件也随之变动。

2. 交换数据

把数据转换为 XML 格式存储不仅可以大大减少交换数据时的复杂性，还可以使这些数据被不同的程序读取，如图 1-2 所示。

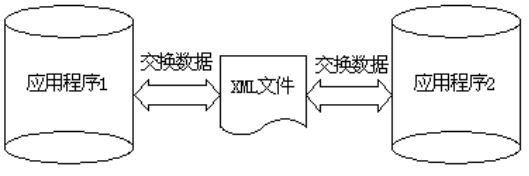

图 1-2　XML 实现不同应用程序之间的数据交互

3. 存储和共享数据

XML 提供了一种与软件和硬件无关的存储和共享数据的方法，大量的数据可以存储到

XML 文件或者数据库中。应用程序可以读写和存储数据，一般的程序可以显示数据。

4. 充分利用数据

XML 是与软、硬件和应用程序无关的，所以可以使数据被更多的用户和设备所利用，而不仅仅是基于 HTML 标准的浏览器。别的客户端和应用程序可以把 XML 文档作为数据源来处理，就像它们对待数据库一样，设计者的数据可以被各种各样的“阅读器”处理。

5. 创建新的语言

利用 XML 可以创建与特定领域有关的标记语言，如 MusicML、MathML、CML、SVG、WML、SMIL 等。XML 允许不同的专业(如音乐、化学、数学等)开发与自己的特定领域有关的标记语言，这就使得该领域的人们可以交换笔记、数据和信息。

XML在数学领域中的应用称为数学标记语言(Mathematical Markup Language，MathML)。MathML适合描述数学方程式，利用它可以把数学公式精确地显示在浏览器上。化学标记语言CML(Chemical Markup Language)可能是第一个XML应用，可以描述分子等信息。

1.2.2　XML 的发展前景

自从 1998 年 2 月发布 XML 1.0 推荐标准后，许多厂商增强了对 XML 的支持力度，包括 Microsoft、IBM、Oracle、Sun 等，它们都相继推出了支持 XML 的产品或改造原有的产品以支持 XML，W3C 也一直在致力于完善 XML 的标准体系。作为互联网的新技术，XML 的应用非常广泛，可以说 XML 已经渗透到了互联网的各个角落。

XML 的开放性、严谨性、灵活性和结构性备受网络开发者的青睐。Web 的飞速发展给予了 XML 充分展示自我的空间，它为使用者提供了更为强大的功能，给程序员带来了更为便利的开发环境。在许多领域，XML 都展现出了卓越的风采。

1. 移动通信领域

随着移动电话与互联网的结合，无线上网的趋势正在形成。有人预言，随着无线带宽的增加和无线上网技术的迅速发展，.move 将代替.com 成为新的潮流。为了满足人们随时随地与互联网连接的需求，Phone.com 联合了 Nokia、Ericsson、Motorola 在 1997 年 6 月建立了 WAP(Wireless Application Protocol，无线应用协议)论坛，旨在利用已有的互联网技术和标准，为移动设备连接互联网建立全球性的统一规范。WAP 是在数字移动电话、因特网或其他个人数字助理(PDA)、计算机应用之间进行通信的全球标准。在 1998 年 5 月，推出了 WAP 规范 1.0 版，WAP 2.0 于 2001 年 8 月正式发布，它在 WAP 1.x 的基础上集成了 Internet 上最新的标准和技术。

WAP 规范包括 WAP 编程模型、无线标记语言(Wireless Markup Language，WML)、微浏览器规范、轻量级协议栈、无线电话应用(WTA)框架、WAP 网关几个组件。其中 WML 是利用 XML 定义的专用于手持设备的置标语言，因 WML 基于 XML，故它较 HTML 更严格。WML 的语法与 XML 一样，它是 XML 的子集。使用 HTML 编写的文件，可以在个人计算机(PC)上用浏览器进行阅读，而使用 WML 编写的文件，则是专用于在手机等一些无线终端显示屏上显示且供人们阅读的。

2. 数据库领域

许多应用程序都使用数据库来管理和存储数据，数据库在数据查询、修改、保存和安全等方面有着其他数据处理手段无法替代的地位。随着网络的迅速发展，让各种应用程序方便地交互各自数据库中的数据显得越来越重要。但不同数据库之间因为数据格式和版本的不同，以及系统设计上的限制，使得它们之间很难快捷、方便地交换数据。

XML 不仅能使应用程序方便地组织数据的结构，而且能帮助各种应用程序方便地交互它们之间的数据。XML 文档可以定义数据结构，代替数据字典，用程序输出建库脚本。应用“元数据模型”技术，对数据源中不同格式的文档数据，可按照预先定义的 XML 模板，以格式说明文档结构统一描述，并提取数据或做进一步处理，最后转换为 XML 格式输出。XML 文件、数据库、网页或文档中的表格，这三者可以互相转换，如图 1-3 所示。

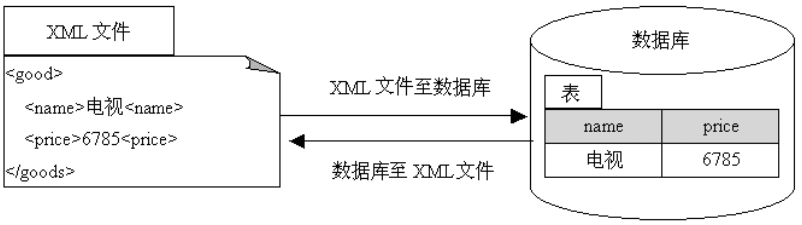

图 1-3　XML 中数据与数据库中记录的相互转换

SOL Server 作为目前比较流行的数据库管理系统，不同的版本都提供了对 XML 的支持。SQL Server 2000 引入了 FORXML 子句(FORXML 是 Select 语句的扩展，返回的查询结果是 XML 流，它以 XML 文档形式形成一个查询结果集)；SQL Server 2005 引入了 XML 数据类型；SQL Server 2008 扩展了合并关系数据库和 XML 数据库解决方案的功能。XML 是新的 Web 数据描述和数据交换的标准数据格式，是数据交换的一种必然趋势，具有非常广阔的应用前景。SQL Server 的不同版本提供对 XML 的增量支持也是趋势使然。

3. 电子商务领域

人类进入 21 世纪以来，互联网、大数据、云计算、人工智能、区块链等现代信息技术正在与实体经济深度融合，我国传统经济逐渐向数字经济转型发展。XML 为网络环境下经济活动主体之间的信息交互提供技术支撑，也为计算机之间的语义交互提供技术支撑。为了提高电子商务的效率，企业内部、企业之间、企业和客户之间进行了广泛的数据集成和应用集成，以实现电子商务交易和业务流程的自动化，而电子商务交易和业务流程自动化的前提是交换数据和业务流程的标准化，这些标准均是采用 XML 定义的。标准是开放的前提，如同有了 TCP/IP 协议就有了开放的互联网世界一样，有了 XML 标准就有了开放的电子商务世界。

电子商务环境下，由于企业的合作伙伴动态多变，需要集成的电子商务数据具有多源异构的特征，因此利用 XML 技术定义数据交换的标准，实现开放式的数据集成尤为重要。电子商务应用集成则普遍采用由 IBM 倡导的 SOA(Service Oriented Architecture，面向服务的架构)，其中服务的描述、发现与集成均采用 XML 描述。另外，基于语义的智能化商务正在迅速发展，描述语义的语言 OWL(Web Ontology Language，万维网本体语言)是基于 XML 的。总之，XML 是电子商务数据集成和应用集成的核心基础技术，没有 XML，就没有电子商务环境下商务过

程的自动化和智能化。

4. 网络出版领域

网络出版，又称互联网出版，是指互联网信息服务提供者将自己创作或他人创作的作品经过选择和编辑加工，上传到互联网或者通过互联网发送到用户端，供公众浏览、阅读、使用或者下载的在线传播行为。

随着互联网的飞速发展，互联网已经成为继报刊、电台、电视台之后的一种新型媒体。在 1998 年 5 月举行的联合国新闻委员会年会上，互联网这一新型媒体被正式冠以“第四媒体”的称号。

网络出版自出现以来，用于信息发布的主要是 HTML 技术，但是这种技术在跨媒体出版时遇到了极大的困难。例如，现在的报纸大多需要同时在网上发布和印刷发行，报社不得不需要两组人力，同时进行印刷组版和网络组版。

另外，随着后 PC 时代的到来，各种如信息家电、手机、PDA 等新的上网设备层出不穷。数字化、网络化已成为主流趋势。便捷化、碎片化的阅读需求对数字化出版提出了更高的要求。基于 XML 的结构化排版，将作品内容和样式分离，具有一次制作、多元多次发布，便于存储和交换等优势，其价值及发展趋势得到了广泛认同。

XML 自出现以来，一直受到业界的广泛关注。虽然由于 XML 的复杂性和灵活性，加上工具的相对缺乏，增加了其使用难度，但毫无疑问，XML 的出现为互联网的发展提供了新的动力，终将成为互联网上全新的开发平台。它促使了新类型的软件和硬件的形成和发展，而这些发展又将反过来促进 XML 的发展。

XML 仍在不断改善，与 XML 相关的技术仍在制定中。XML 需要强大的新工具在文档中显示丰富、复杂的数据，XML 会对终端用户在网上的行为不断进行改进，这有助于许多商业应用的实现。XML 作为一个数据标准，将会开发互联网上的众多新用途。

1.3 XML 相关技术

XML 并不仅仅包括 XML 标记语言，它还包括很多相关的规范，如文档模式技术、文档显示技术、文档查询技术、文档解析技术、文档链接技术及文档定位技术等。基于 XML 的这些规范，还有很多高层的应用协议，如 SOAP(Simple Object Access Protocol)和 BizTalk 等。下面将对其中比较关键的几种技术进行简单介绍。

1. 文档模式技术

XML 文档为了保证数据交换的准确性，需要满足语义规范。DTD(Document Type Definition，文档类型定义)是 W3C 推荐的验证 XML 文档的正式规范。也就是说，一个实用的 XML 文档要符合 DTD 的语法规定，这样既能保证 XML 文档的易读性，又能充分体现数据信息之间的关系，从而能够更好地描述数据。

DTD 是用于描述、约束 XML 文档结构的一种方法。它规定了文档的逻辑结构，可以定义文档的语法，而文档的语法反过来能够让 XML 语法分析程序确认某个页面标记使用的合法性。DTD 定义页面的元素、元素的属性以及元素和属性之间的关系。DTD 文件是 XML 文件的类型

定义文件，相当于 XML 文件的法律性文件，如果一个 XML 文件不满足其关联的 DTD 文件的约束，就不是一个有效的 XML 文件。

DTD 不是强制性的。对于简单应用程序来说，开发人员不需要建立他们自己的 DTD，可以使用预先定义的公共 DTD，或者根本就不使用。

下面是一个简单的 DTD 文档。

```
<!ELEMENT persons (person*)>
<!ELEMENT person (name,sex,birthday)>
<!ELEMENT name (#PCDATA)>
<!ELEMENT sex (#PCDATA)>
<!ELEMENT birthday (#PCDATA)>
```

DTD 本身是专门为 SGML 的确认规则开发的，它并不符合 XML 规范，而且语法复杂，难以掌握。由于 DTD 存在着种种缺陷，促使 W3C 组织致力于寻求一种新的机制来取代它。在众多标准中，微软公司在 2000 年发布的 XML Schema 工作草案引人注目，它具有完全符合 XML 语法、丰富的数据类型、良好的可扩展性以及易于处理等优点。Schema 不仅能实现 DTD 的功能，还能定义文本数据的实际意义。Schema 文件是 XML 文件的模式定义文件。下面是一个简单的 Schema 文件。

```
<?xml version="1.0"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
<xsd:element name="persons">
<xsd:complexType>
      <xsd:sequence>
      <xsd:element name="person" minOccurs="0" maxOccurs="unbounded">
      <!--设置 person 的子标记-->
            <xsd:complexType>
            <xsd:sequence>
                  <xsd:element name="name" type="xsd:string"/>
                  <xsd:element name="sex" type="xsd:string"/>
                  <xsd:element name="birthday" type="xsd:string"/>
          </xsd:sequence>
         </xsd:complexType>
    </xsd:element>
    </xsd:sequence>
</xsd:complexType>
</xsd:element>
</xsd:schema>
```

2. 文档显示技术

XML 是内容(数据)和显示格式相分离的语言，其特点就是数据与样式的分离，不提供数据的显示功能，它的显示功能由称为样式表的相关技术来完成，这样就可以按照用户的意愿为同一数据任意添加多种样式，如图 1-4 所示。

使用独立的样式表文件制定显示格式的优势在于：对同一份数据文件可以制定出不同的样式风格，这些不同的样式可以应用于不同的场合，使数据能够更合理、更有针对性地表现出来，从而提高了数据的重用性。

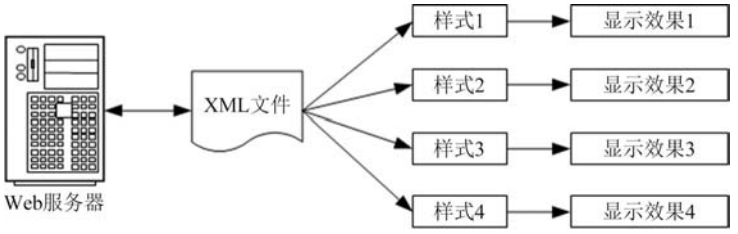

图 1-4　一种数据多种样式

W3C 提供了两种通用的样式语言，即 CSS(Cascading Style Sheet，层叠样式表)和 XSL(eXtensible Style Language，可扩展样式语言)。

其中，CSS 是随着 HTML 的出现而产生的，用于设置字体样式等内容，CSS 就是一组规则的集合。CSS 可以控制 XML 文档的显示，但不会改变源文档的结构。而 XSL 是专门为 XML 设计的，是一种特殊的 XML 文件，不仅能用来显示 XML 文档，还可以把一个 XML 文档转换为另一个 XML 文档。

【例 1-2】使用 CSS 文件和 XSL 文件显示 example1-2.xml 文件的内容。

example1-2.xml 文件的源代码如下。

```
<?xml version="1.0"    encoding="gb2312"?>
<?xml-stylesheet type="text/css"    href="show.css" ?>
<persons>
    <person>
        <name>小李</name>
        <sex>male</sex>
        <birthday>1981.12.25</birthday>
    </person>
    <person>
        <name>小陈</name>
        <sex>female</sex>
        <birthday>1974.10.20</birthday>
    </person>
 </persons>
```

如果使用 CSS 显示 XML 数据内容，则 CSS 文件 show.css 的代码如下。

```
name{ display:block; font-size:18px;}
sex{ display:block;font-size:18px;}
birthday{ display:block; font-size:18px;}
```

显示效果如图 1-5 所示。

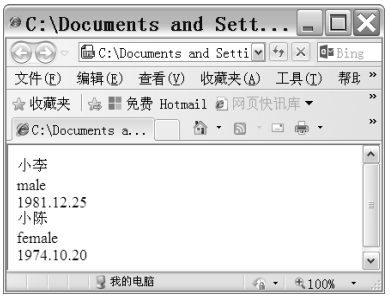

图 1-5　使用 CSS 文件 show.css 显示 example1-2.xml 文件的内容

使用 XSL 文件 show.xslt 显示同一 XML 文件内容，XSL 文件的代码如下。

```
<?xml version="1.0" ?>
<xsl:stylesheet  version="1.0"  xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:template match="/">
<html>
<body>
<center>
  <table border="1">
   <tr><td>name</td> <td>sex</td><td>birthday</td></tr>
   <xsl:for-each select="persons/person">
   <tr>
       <td><xsl:value-of select="name" /></td>
       <td><xsl:value-of select="sex" /></td>
      <td><xsl:value-of select="birthday" /><br /></td>
  </tr>
  </xsl:for-each>
 </table>
</center>
</body>
</html>
</xsl:template>
</xsl:stylesheet>
```

注意：要对【例 1-2】example1-2.xml 中的第二行代码进行如下修改。

```
<?xml-stylesheet  type="text/xsl"  href="show.xslt" ?>
```

其完整的代码在 example1-3.xml 中，这时 XML 文档中的数据以表格的形式显示，如图 1-6 所示。

图 1-6　使用 XSL 文件 show.xslt 显示 example1-3.xml 文件的内容

3. 文档解析技术

为了有效地使用 XML，必须通过编程来访问数据。XML 解析器是 XML 文档和应用程序之间存在的一个软件组织，主要起桥梁的作用，为应用程序从 XML 中提取所需要的数据。XML 解析器最基本的功能就是检查文档格式是否良好，大多数解析器还能够判断文档是否符合

DTD/Schema 规范。

XML 解析器分成两大类：综合解析器和专用解析器。综合解析器除了具有分析 XML 文件代码语法的功能外，还具有其他功能，如解析出需要的数据等。IE 6.0 就是一个综合解析器。专用解析器就是一个应用程序，是为了某一特定功能而设计的，只能分析出一段 XML 程序是否合法等，如微软的 Internet Explorer 浏览器就内置了 MSXML 解析器。综合解析器又分为基于 DOM 的解析器和基于事件的解析器。

DOM 解析器的核心是在内存中建立一个和 XML 文件相对应的树结构，会占用很多内存空间，适用于解析小型的 XML 文件。

基于事件的解析器，如简单应用程序接口(SAX)在解析的过程中，并不在内存中建立这样的一个树结构。它的核心是事件处理机制，会把 XML 文件转换成事件流的形式传递给解析器的处理器，处理器逐个地对每个事件进行处理。所以，基于事件的解析器占用很少的内存，具有更高的工作效率，可以解析大型的 XML 文件。

DOM 是由 W3C 推荐的处理 XML 文档的规范，而 SAX 并不是 W3C 推荐的标准，但却是整个 XML 行业的事实规范。

4. 文档链接技术

Web 迅速发展和普及的一个重要因素是 HTML 的应用，而 HTML 真正的强大之处在于它可在文档中嵌入超链接。超链接是描述 HTML 文档中不同部分之间关系的一种技术，XML 的链接功能比 HTML 更强大，在 XML 中，超链接被扩充为独立的链接语言。XML 的链接技术分为两部分：XLink(XML Linking Language)和 XPointer(XML Pointer Language)。XLink 定义一个文档如何与另一个文档链接(类似 HTML 中的外部链接)，而 XPointer 则规定了 XML 文档中不同位置之间的链接规范(类似 HTML 中的内部链接)。

XLink 的目的是描述 Internet 上任一页面上的任何一部分和 Internet 上其他页面上的某些部分之间的关系。XLink 的一个重要应用是用于超文本链接。简单的超文本类似于 HTML 中的超链接标记<a>，但 XLink 中定义的链接远远超出了目前使用的 HTML 链接。XLink 可以有多个链接终点，可以从不同的方向进行遍历，还可以将链接存储于独立于引用文档的数据库中。

在 XLink 中，并不涉及标识不同类型数据位置的方法，XLink 依赖于不同的机制来标识想要链接的资源(如统一资源标识符)。因此，W3C 推出了 XPointer，用于构造 XML 文档的内部结构。XPointer 可以链接到一个具体的对象上，这个对象可以是一个网页、网页的一部分、网页中的一个元素，甚至网页中某行的某几个字。XPointer 是对 XPath 概念和寻址方法的扩展，可以直接在 URL 中对 XML 文件的不同部分进行寻址，为 XML 的超链接提供基本条件。

5. 文档查询技术

W3C 推荐的 XML 的查询语言是 XQuery，其全称是 XML Query。XQuery 是一门用于查询 XML 数据的新语言，它由 W3C 的 XML 查询工作组设计，是用于查询 XML 数据的查询语言。类似于 SQL 用于查询关系数据库，XQuery 用于查询 XML 数据。

XQuery 查询的 XML 数据不仅可以是 XML 文档，还可以是任何能以 XML 形式呈现的数据，包括数据库。从这个意义上讲，XQuery 可以非常方便地从 XML 数据中提取出应用程序所需的数据。

6. 文档定位技术

在转换 XML 文档时，可能需要处理其中的一部分数据。那么，如何查找和定位 XML 文档中的数据呢？XML Path Language(XPath)是一种用于对 XML 文档各部分进行定位的语言，用于在 XML 文档中查找信息。XPath 可在 XML 文件中快速找到某个特定的标记，可用于在 XML 文档中对元素和属性进行遍历。

其他 XML 程序可利用 XPath 在 XML 文档中对元素和属性进行导航，它主要用于为 XSLT、XPointer 以及其他 XML 技术提供服务。XSLT、XPointer 等技术需要依赖于 XPath 来定位 XML 文档中的元素和属性等节点。

XPath 和 XQuery 在某些方面很相似。XPath 还是 XQuery 不可分割的一部分。这两种语言都能够从 XM 文档或者 XML 文档存储库中选择数据。虽然 XPath 和 XQuery 都能实现一些相同的功能，但是 XPath 比较简洁而 XQuery 更加强大和灵活。对于很多查询来说，使用 XPath 非常合适。例如，若要通过 XML 文档中的部分记录建立电话号码的无序列表，则使用 XPath 实现最简单。但是若要表达更复杂的记录选择条件、转换结果集或者进行递归查询，则使用 XQuery 更为合适。

1.4 XML 编辑工具

XML 只是一种简单的文本文件，其扩展名为.xml，因此开发者完全可以使用普通的文本工具来编辑 XML 文档。当然，选择一款专业的 XML 编辑工具则会起到事半功倍的作用。

1.4.1 普通文本编辑工具

下面以“记事本”为例说明编写 XML 文件的过程。

(1) 编辑 XML 文件。单击“文件”菜单下的“新建”命令，在“记事本”中输入【例 1-1】example1-1.xml 的代码。

(2) 保存 XML 文件。在“文件”菜单下选择“另存为”命令，以文件名 example1-1.xml 保存该文件，保存类型为所有文件，编码为 ANSI，如图 1-7 所示。

图 1-7 在“记事本”中保存 XML 文件

注意：

如果 XML 文件指定了文件的编码，则在保存时也必须使用同样的编码，这样 XML 解析器才能识别 XML 中的标记并能正确地解析出所标记的内容。例如，在编写 XML 文件时指定文件的编码为 UTF-8，则保存文件时编码也应选为 UTF-8。

(3) 查看 XML 文件。XML 文件一般是配合其他应用程序而使用的，要想单独运行 XML 文件，最简单的方法就是用 IE 浏览器直接打开 XML 文件。在浏览器中打开 example1-1.xml，浏览器将显示该文件的内容，如图 1-8 所示。

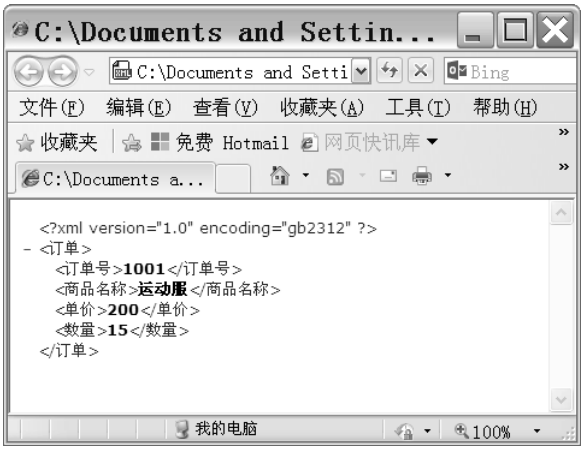

图 1-8　在 IE 中显示 XML 文档

1.4.2　本书的开发环境

开发环境集成了代码的编写和解析等功能，方便了用户应用和开发。XML 的开发应用环境包括 XML 编辑工具、验证工具、解析工具和浏览工具 4 项。目前市面上单项功能的工具和多项功能的工具都有很多，如 XMLWriter、XML Spy、Stylus Studio、Visual XML 等。由于目前使用 XML Spy 的用户较多，因此本书选用 XML Spy 2013 作为 XML 的开发应用环境。

1.4.3　XML Spy 简介

Altova GmbH 公司的 XML Spy 是处理 XML 的一整套工具。它支持以“所见即所得”的方式来编辑 XML 文件，支持 Unicode、UTF-8 等多种字符集。不仅如此，它还支持创建 Java Web、EJB 等组件的配置描述文件。

XML Spy 支持对 XML 进行验证，它支持验证 Well-formed(格式良好的)和 Validated(有效的)两种类型的 XML 文档。XML Spy 对 DTD、Schema 等语义约束工具提供了良好的支持，既可为现有的 XML 文档生成对应的 DTD 或 Schema 语义约束文档，又可根据 DTD 或 Schema 生成 XML 文档结构。

XML Spy 还提供了强有力的样式表设计，既可编写 CSS，又可编写 XSLT 等样式表，对 XSLT 1.0、XSLT 2.0 和 XSLT 3.0 都提供了良好的支持，并集成了功能强大的 XSLT 调试工具。

XML Spy 对 XQuery 也提供了强大的支持，包括集成的 XQuery 环境，并允许直接浏览 XQuery 的查询结果。

1.4.4 使用 XML Spy 编辑 XML 文档

使用 XML Spy 编辑 XML 文档之前，应先下载和安装 XML Spy。目前，提供 XML Spy 开发环境下载的网站很多，也可以从 Altova 的官方网站(http://www.altova.com/)获取最新的试用版本。安装 XML Spy 和安装普通程序没有任何区别，此处不再赘述。

安装好 XML Spy 之后，就可以用它创建文档了。该创建过程可分为新建文档、添加内容、验证和保存 4 个步骤。

1. 新建文档

新建文档的操作步骤如下。

(1) 双击 Altova XML Spy 的快捷方式，启动 XML Spy 2013，界面如图 1-9 所示。

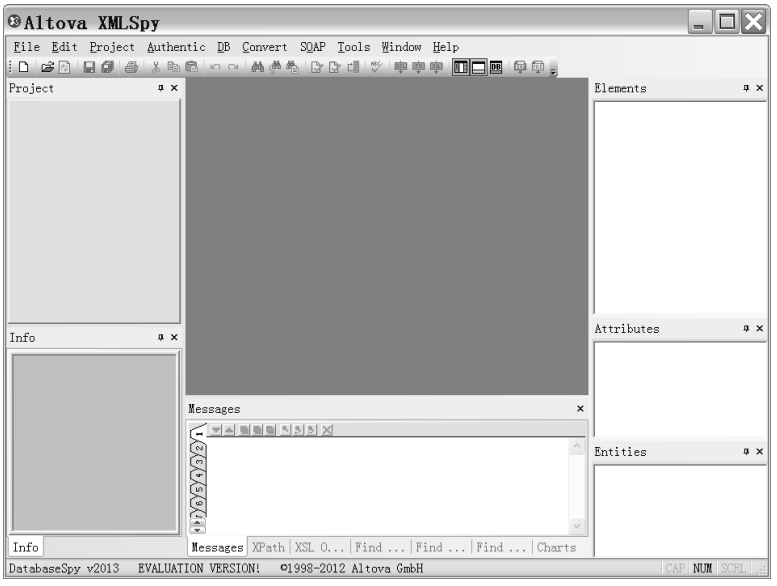

图 1-9 XML Spy 2013 的界面

(2) 单击 File | New 命令，弹出 Create new document 对话框，如图 1-10 所示。

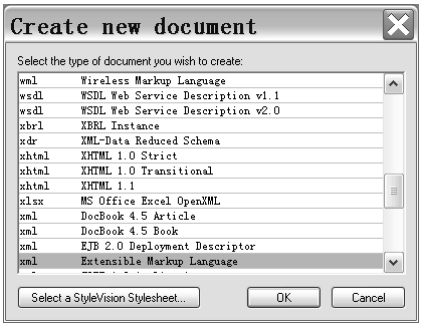

图 1-10 Create new document 对话框

(3) 选择 xml Extensible Markup Language 选项，单击 OK 按钮，弹出如图 1-11 所示的提示，询问所创建的 XML 文档是基于 DTD 还是基于 Schema。

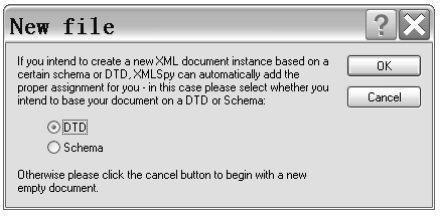

图 1-11　新建文档

(4) 单击 Cancel 按钮，创建一个无文档类型说明的 XML 文档。这样，一个空白的 XML 文档即创建完毕，如图 1-12 所示。

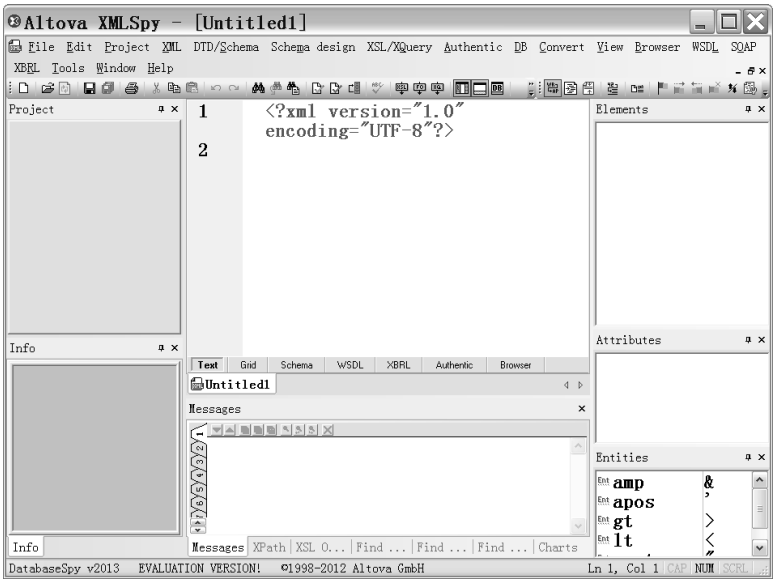

图 1-12　创建的空白 XML 文档

2. 添加内容

新建文档后，就可以添加所需的数据了。例如，可以将例 1-1 中 example1-1.xml 的代码输入文本框中，添加内容后的 XML 文档界面如图 1-13 所示。

图 1-13　添加内容

3. 验证

由于上面的 XML 文档并未指定 DTD 或 Schema，因此只可能是格式良好的文档，可以用 XML Spy 验证文档的格式是否良好，操作步骤如下。

(1) 添加内容后，单击 XML | Check well-formedness 命令，对 XML 文档进行结构完整性检测，如图 1-14 所示。

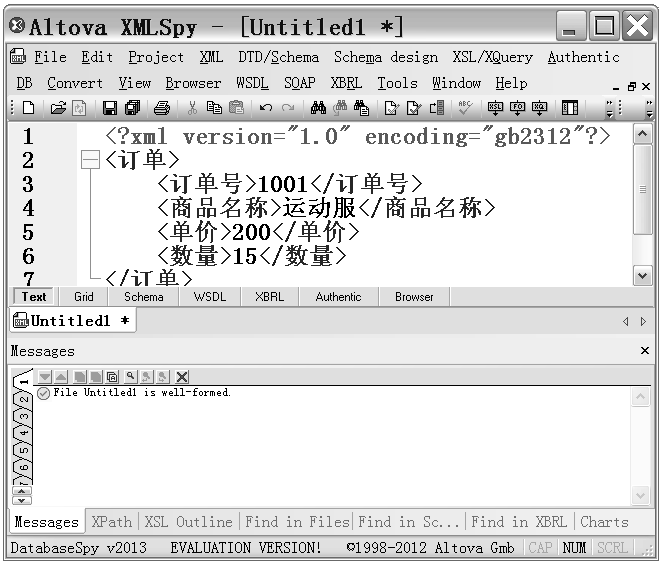

图 1-14　结构完整性检测

(2) 单击 OK 按钮，再单击右下方的 Browser 按钮，在 XML Spy 中的集成浏览界面进行显示，如图 1-15 所示。

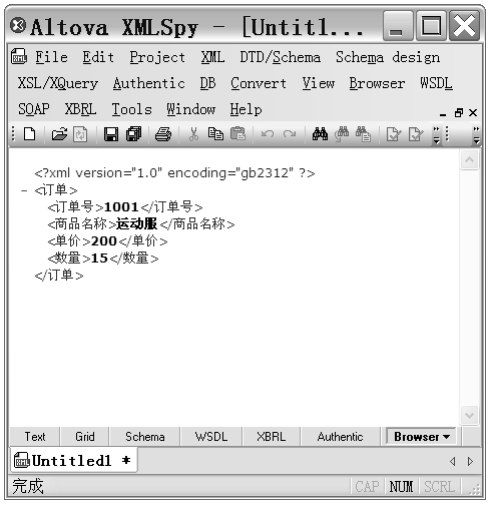

图 1-15　在集成浏览界面中显示文档

4. 保存

在结构完整性检测和集成浏览显示均正常的情况下，单击 File | Save 命令，将其保存。

1.4.5　XML Spy 的视图格式

XML Spy 不仅提供了用于编辑和创建 XML 文档的功能，还提供了多个编辑视图以便选择，通过单击文档显示区域下面的不同标签，可以在不同的视图之间切换，如图 1-16 所示。

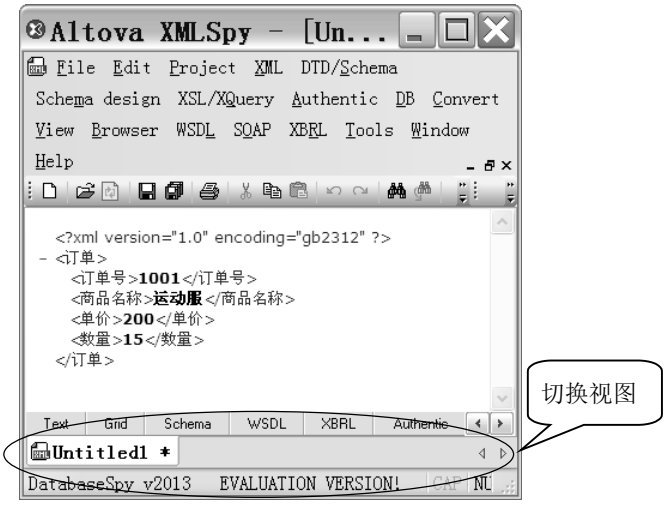

图 1-16　以 Text 视图显示 XML 文档

- Text：一种最基本的编辑视图。可以查看和修改文档的源代码，并以不同的颜色标注不同的元素。
- Grid：一种包含层次结构的编辑视图，它用一系列嵌套栅格展示了 XML 文档的逻辑结构。
- Schema/WSDL：仅在使用 XML 模式或 WSDL 文档时可用。
- Authentic：XMLSpy 特有的视图，它使用 StyleVision 样式表(stylesheet)来显示 XML。StyleVision 样式表是 XML 文件的一个图形化覆盖图。它包括控制、验证和图表等，让那些不习惯处理尖括号的人使用 XML 时更轻松。
- Browser：使用 Internet Explorer 来显示 XML 文档(需要 IE 5 或以上的版本)，支持 CSS 和 XSL。

1.5　本章小结

本章简要介绍了 XML 的发展历史和特点。起初，XML 的诞生旨在更好地进行数据交换。它基于 SGML 发展起来，是 SGML 的一个精简子集。另外，它也是一种元标记语言，具备自我解释性，可用于编写其他新的标记语言。本章首先讲解了标记语言的发展历史、XML 的发展历史和特点、XML 与 HTML 的区别，并对 XML 的应用现状和发展前景进行了简要描述。其次，本章介绍了 XML Spy 开发环境，并概述了如何在该环境下创建 XML 文档。

1.6 思考和练习

1. 如何理解 XML？XML 的特点是什么？
2. 和 XML 有关的技术有哪些？
3. 请分析 HTML 与 XML 的异同点。
4. XML 有哪些用途？
5. 编写一个简单的 XML 文件，并用 IE 查看其运行效果。