

# 基于 XML 的通用数据交换系统<sup>①</sup>

洪欣, 陈维斌

(华侨大学 计算机科学与技术学院, 泉州 362021)

**摘要:** 尽管当前许多数据库都提供对 XML 格式数据的支持, 但是由于接口不同, 缺乏通用性。采用 java 技术, 开发了一个与平台无关的通用数据交换系统, 通过该系统的基于 XML 文档的虚拟关系数据库管理平台 and 基于关系数据库的虚拟 XML 数据库管理平台, 实现平台无关的数据交换。

**关键词:** XML; 关系; 数据映射; 数据交换; 平台无关

## Common Data Exchange System Based on XML

HONG Xin, CHEN Wei-Bin

(College of Computer Science and Technology, Huaqiao University, Quanzhou 362021, China)

**Abstract:** Although many databases support the XML format data, the different interfaces they apply lack universality. A platform independent Common Data Exchange System is established with java technology. Through the Platform of Virtual Relational Database Manage System and the Platform of Virtual XML Database Manage System, the XML-document and the Relational-data can be exchanged by ignoring the platform.

**Key words:** XML; relation; data mapping; data exchange; platform independence

XML(Extensible Markup Language)即可扩展标记语言, 它与 HTML 一样, 都是 SGML(Standard Generalized Markup Language,标准通用标记语言)。XML 是 Internet 环境中跨平台的, 依赖于内容的技术, 是当前处理结构化文档信息的有力工具<sup>[1]</sup>。XML 的自描述性使其成为网络中被广泛使用的数据描述语言。本课题开发了一个平台无关的数据交换系统, 实现 XML 文档的数据与关系数据库的数据的双向传送<sup>[2-4]</sup>。

### 1 各数据库应用系统对XML的支持

由于 XML 在网络中的广泛应用, 许多流行的数据库应用平台纷纷提供对 XML 的支持。SQL Server 2008 提供了 FOR XML 和 OPEN XML 的 XML 访问接口(图 1), Oracle 提供了对 Java ,C 和 C++的 XML 解析器<sup>[5-7]</sup>(图 2)。

由以上分析可知, 虽然许多大型数据库提供了对 XML 数据的访问支持, 但是这种支持依赖于原有的数

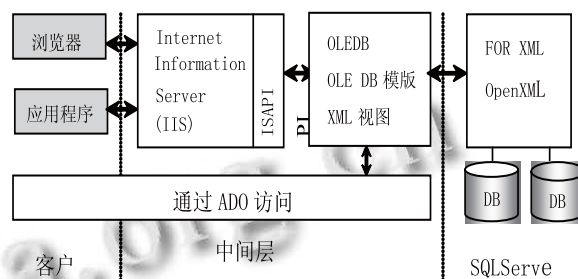


图 1 XML 访问 SQL Server 的体系结构概览

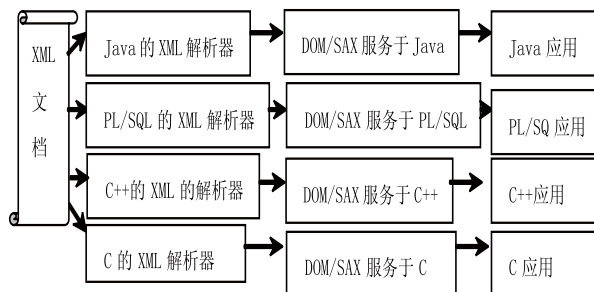


图 2 Oracle 提供了 Java,C 和 C++的 XML 解析器

① 基金项目:福建省自然科学基金(2009J01288);福建省青年人才项目(2007F3062);华侨大学校科研基金(08HZR18);泉州市科技计划 (2009G5.2010G4) 收稿时间:2011-04-19;收到修改稿时间:2011-05-21

数据库系统,不同的数据库系统对于 XML 格式数据的支持接口不同,缺乏通用性,对于异构数据库中对于 XML 数据的存取和处理造成极大困难。因此,有必要开发一个与平台无关的 XML 与关系数据库的数据交换平台。

## 2 系统设计

系统采用 Java 语言,提供了两个虚拟平台:基于 XML 的虚拟关系数据库平台和基于关系数据库的虚拟 XML 平台;实现 XML 文档和关系数据库的相互映射。系统的结构图如图 3 所示。

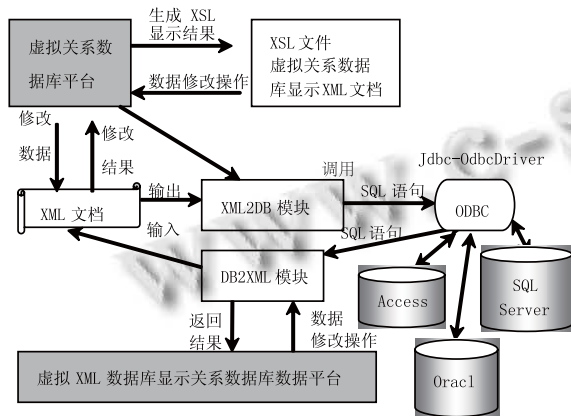


图 3 XML 文档与关系数据库的双向转换系统

### 2.1 基于 XML 的虚拟关系数据库平台

为 XML 文档提供一个虚拟的关系数据库平台,使得 XML 文档在没有关系数据库支持的情况下,能够以表格得形式显示 XML 文档的内容,在该虚拟的关系数据库平台上,能够对数据进行删除添加等操作,仿佛在一个关系数据库上对数据进行处理。由该虚拟平台对数据的修改操作是通过一个“虚拟关系数据库模块”进行,并以 XSL 的方式显示 XML 文档内容,由于浏览器的支持,所生成的 XSL 文件可以直接用于 XML 数据的网上发布。

### 2.2 基于关系数据库的虚拟 XML 平台

为关系数据库提供了一个虚拟的 XML 数据库平台,使得关系数据库中的数据能够以 XML 文档的层次形式显示,在该虚拟平台上能够对数据进行删除添加等操作,仿佛在一个 XML 数据库上对数据进行处理。

## 3 系统工作原理

### 3.1 XML 文档映射为关系数据库 XML2DB 子系统

在将 XML 数据转换为关系数据的方向上,对 XML

文档进行处理,生成相应的 SQL 语句,通过 ODBC 传送给关系数据库,生成关系表,同时可以生成 XSL 样式表,使用该样式表可以绕过关系数据库管理系统,直接将源 XML 文档在浏览器中以关系表的形式显示,相当于一个虚拟的关系数据库。

由于模式在 XML 文档中不是必须的,因此将 XML 文档映射为关系数据库分为不带 XDR 模式的映射,带有 XDR 模式的映射两种。为了程序的通用性,本实验系统对不带 XDR 模式的 XML 文档的映射包含三个步骤:首先生成 XDR 模式;然后应用含有 XDR 模式的映射方法,对 XML 文档进行映射生成关系表和 XSL 样式表;最后通过填入数据模块将 XML 文档的数据映射为关系数据库的数据,系统工作流程如图 4 所示<sup>[8-12]</sup>。

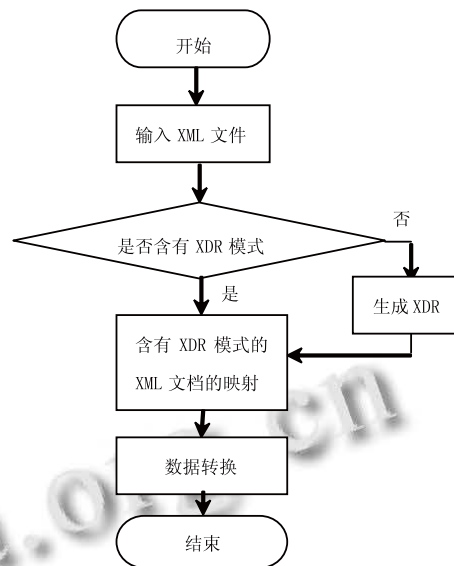


图 4 XML 文档到关系数据库的映射系统工作流程

### 3.2 关系数据到 XML 数据的映射 DB2XML 子系统

在将关系数据库的数据转换为 XML 数据的方向上,也是通过 ODBC 发送 SQL 语句,从关系数据库获取数据然后经过处理,生成 XDR 模式及 XML 文档,同时通过向 ODBC 接口发送 SQL 语句,以 XML 的层次嵌套形式显示关系数据库中的关系表关系及数据,这样关系数据库就仿佛一个虚拟的 XML 数据库。由于系统采用的是 ODBC 接口,因此可以与任何关系数据库实现交互,具有通用性、可移植性和平台无关性。由于系统通过 SQL 语句对数据进行处理,因此在对存储在关系数据库中的 XML 数据的处理就像关系数据

一样方便,能够实现对数据的所有操作处理。另一方面在对 XML 文档的数据的修改处理,只要对文件处理后重写即可,较为简单。

关系数据到 XML 数据的映射分为两种,一种是只输出单表的数据,另一种是将表及关系一同输出。第一种情况比较简单,不考虑外码所指的内容,只输出本表的数据即可。第二种情况,需要考虑表间的链接,取得外码所指的表的数据,能够输出该表数据及其所有子表的数据。系统工作流程如图 5 所示。

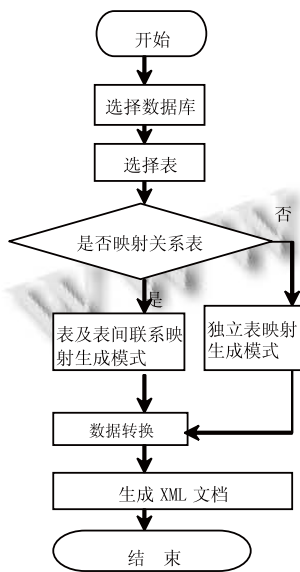


图 5 关系数据到 XML 文档的映射的系统工作流程

### 3 结语

为了提高对 XML 数据的存取效率,本课题没有采用 Oracle 系统所采用的存取方式,而是将 XML 数据存储的关系表中,与 SQL SERVER 2008 采用的方式相同,但是比 SQL SERVER 2008 所提供的接口有所改进,可以通过对 XML 文档的结构映射创建关系表。并且采用了较之于 SQL SERVER 2008 更为高效的映射方法,既产生较少的关系表,最大限度地保存了 XML 中与数据无关的控制参数,又保证这些控制参数的存储方式不至于影响查询效率。

在输出处理方面,当前的数据库系统所提供的对 XML 格式的支持,仅限于将数据库与 XML 文档的相

互转换,却没有提供所输出的 XML 文档的显示方式,在浏览器中只能看到一个层次结构的 XML 文档,XML 数据的显示并不直观。在本课题的实验系统中,提供了基于 XML 的虚拟关系数据库,在从关系数据库转换并输出 XML 文档的同时,输出相应的 XSL 样式表,使得输出的 XML 文档能够在浏览器中以关系表的形式显示。

在关系数据库的方面,为了直观显示关系数据库中表格的关系,本系统提供了一个虚拟的 XML 数据显示平台,使得关系表表间的联系能够以 XML 文档的层次方式显示,解决了 XML 文档映射到关系数据库之后,表间关系的显示不够直观的问题。

### 参考文献

- 1 xml. <http://baike.baidu.com/view/63.htm>,2009-5-12.
- 2 康晓兵,张二虎,吴学毅.一种 XML Schema 模式到关系模式的映射算法.计算机应用,2004,24(5):106-108.
- 3 王志平.基于 XML 的异质多数据库集成系统的设计与实现.河南大学学报(自然科学版),2007,37(5):530-532.
- 4 洪欣.基于 XDR 模式的 XML 文档与关系数据库的映射技术研究.泉州:华侨大学,2004.
- 5 Burke PJ.方力,等译.SQL Server 2000XML 高级编程.北京:中国电力出版社,2003.
- 6 (美)马尔科姆.莱恩工作室译. Microsoft SQL Server2000 与 XML 程序设计.北京:北京大学出版社,2001.
- 7 胡欣杰.Oracle 9i XML 网络数据库开发指南.北京:北京希望电子出版社,2002.
- 8 洪欣.基于 XDR 模式的 XML 文档与关系数据库的映射技术研究.泉州:华侨大学,2004.
- 9 洪欣,陈维斌,蹇崇军.XML 模型到关系模型的映射.华侨大学学报(自然科学版),2009,30(6):83-87.
- 10 洪欣,陈维斌,蹇崇军.XML 数据到关系数据的映射.华侨大学学报(自然科学版),2011.1.
- 11 洪欣,陈维斌,杜吉祥.抽取 XML 模式到关系模式.华侨大学学报,2010,31(3):288-291.
- 12 洪欣,陈维斌.XML 模式与关系模式的无损映射分析.山东理工大学学报,2010,24(5):23-30