

基于XML的电子商务数据模式的转换研究

陈雪刚¹,程杰仁^{1,2}

(1. 湘南学院 数学系, 湖南 郴州 423000; 2. 国防科技大学 计算机学院, 湖南 长沙 410073)

摘要: XML已成为数据交换的事实标准, XML在电子商务中的应用越来越广泛。为了更好地利用关系数据库, 需要将关系数据转换为XML数据。文中对Schema和DTD进行了比较, Schema模式是XML的首选模式, 并分析了Schema的结构和作用, 结合电子商务开发提出了关系模式到XML Schema的转换。实现了关系模式到XML Schema的完全转换, 该转换能更有效地等价描述关系模式, 保证了模式转换后数据的一致性和完整性。

关键词: 关系模式; XML; XML Schema; 电子商务

中图分类号: TP311

文献标识码: A

文章编号: 1673-629X(2008)09-0202-03

Research of Transforming for Model of Electronic Commerce Data Based on XML

CHEN Xue-gang¹, CHENG Jie-ren^{1,2}

(1. Department of Mathematics, Xiangnan University, Chenzhou 423000, China;

2. School of Computer Science, National University of Defense Technology, Changsha 410073, China)

Abstract: XML has become the standard for data exchange in fact. XML is used widely in E-commerce applications, and relational data is converted to XML data, in order to make better use of relational database. XML Schema was compared with DTD, and the structure and the role of XML Schema are analyzed. Bring about a constraint algorithm of converting from relational schema to XML Schema with the development of E-commerce. And turning relational model into XML schema is achieved completely, and the algorithm can be more effective described relational model equivalent, ensuring integrity and consistency of data after transforming of model.

Key words: relational schema; XML; XML Schema; electronic commerce

0 引言

随着网络的迅速发展,数据的日益增多,由于不同企业使用不同的数据格式,为了统一数据格式,XML^[1](eXtensible Markup Language)应运而生,XML语言是为文档交换所设计,以一种开放的、自我描述方式定义数据结构,既描述数据更突出结构的描述,实现了跨平台数据交换和数据传输,如今XML语言作为一种新的数据交换标准在Internet上得到了快速发展,XML的应用越来越广泛,尤其在电子商务领域中有更好的应用前途,是解决电子商务领域问题的关键技术,XML与后台关系数据库的相互转化成为研究热点。2001年5月W3C(万维网联盟)正式发布XML Schema的推荐标准^[2],是DTD(文件类型定义)之后第

二代用来描述XML文件的标准。由于XML Schema优越性,文中所研究的是关系模式到XML Schema的转换问题,该转换用于电子商务(Electronic Commerce)中后台关系数据转换到XML文档是等价的,符合第三范式,确保了数据转换前后的一致性和完整性。

1 XML Schema

1.1 XML Schema与DTD比较

(1)易用性:XML Schema^[3]是按标准XML规则编写的,容易掌握;DTD是用不同于XML的语言编写的,需要不同的分析器技术。

(2)易操作:XML Schema更易于应用程序处理元数据。

(3)扩展性:XML Schema提供namespace的支持,能使置标唯一,DTD则不能。

(4)互换性:XML Schema丰富的内置数据类型,并支持用户对数据类型的扩展,能胜任关系模式在数据上的描述的需求,XMLDTD在对关系数据的描述方

收稿日期:2007-12-04

基金项目:湖南省教育科研项目资助(07C718);湘南学院资助科研项目(06Y017)

作者简介:陈雪刚(1976-),男,湖南衡南人,讲师,研究方向为XML数据库、电子商务、网络安全和软件开发。

面明显存在着不足,它有限的数据类型根本无法完成对关系数据数据类型的——映射,对强制性结构化更是无计可施。

(5)继承性:XML Schema 支持软件技术领域中的主流概念——继承,DTD 不支持数据的继承。

(6)规范性:XML Schema 不仅可以定义 XML 文档的结构而且还允许约束文档的内容,提供了一系列组件严格约束模式结构和内容。

1.2 XML Schema 结构

元素和属性是创建 XML 文档的主要构建材料。在 W3C XML Schema 中,元素通过使用<element>元素实现。在 XSDL 中,属性实现的方法是使用<attribute>元素。W3C XML Schema 可以把 XML 文档中的元素和属性声明为特定的类型,准许解析器检验文档的内容和结构。XML Schema 定义了两类主要的数据类型:预定义简单类型和复杂类型,以及提供组合器来控制结构。

(1)简单类型。

- * 原子类型具有不可分割的值;
- * 列表类型的值为用空白符隔开的原子值列表;
- * 联合类型的值可以是原子值,也可是列表值。

(2)复杂类型。

复杂类型的内容类型有 4 种:简单类型、纯元素类型、混合类型和空类型。

(3)利用组合器控制结构。

模式组允许把子元素声明或引用组合起来,从而构建更有意义的内容模型。模式组共有 3 种:all、choice、sequence。

1.3 XML Schema 的作用

Schema 文档就是用来验证 XML 文档的正确性,用来判断实例是否符合模式中所描述的所有约束^[4]。涉及到检查实例中所有的元素和属性。

Schema 主要作用如下:

- (1) 约束数据和数据结构。
- (2) 保证 XML 文档中出现的元素和属性都是模式规定的词汇表。
- (3) 确保一致性。
- (4) 为元素和属性添加默认值和固定值。
- (5) 扩充实例。

2 关系模式到 XML 模式的转换

2.1 基于 XML 的电子商务数据转换结构

根据目前大多数企业的电子商务平台的建立都是基于关系数据库的后台系统,随着 XML 的应用,XML 已成为 Internet 的数据交换标准,企业的子公司或客

户端根据需从服务器搜索所需要的信息,把查询到的关系数据转换为 XML 文档,再通过 XML 技术,如 XSLT(Extensible Stylesheet Language Transformation,可扩展的样式表语言)将 XML 数据格式转换成 HTML 描述格式或者其它数据格式。基于 XML 的电子商务数据转换结构如图 1 所示。

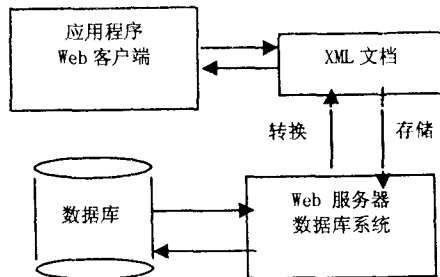


图1 基于XML的电子商务数据转换结构

2.2 转换过程

关系数据库是由许多关系表组成,关系表是由相应属性组成,在开发电子商务中有这样的关系数据库,销售商所售的书和消费者,他们的关系表如下:

book(book_ID,book_name,writer)

customer(customer_name,book_ID,customer_address)

其中,book 关系表是书的相关信息,book_ID 是其主关键字;customer 关系表是购书者的信息,主关键字是 customer_name,而 book_ID 是表 book 的主关键字,是外关键字,关系数据库是基于此,表之间产生关联,进行数据表示、约束和操作。现将关系模式转换到 XML Schema 步骤如下:

(1) 关系数据库是由许多关系表构成,假设 book 关系表和 customer 关系表组成了库。如图 2 所示。

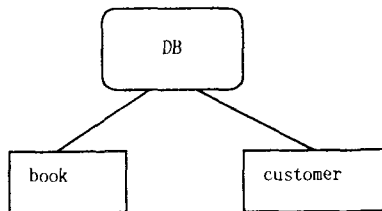


图2 关系数据库的层次结构

在转换时,book 和 customer 作为 DB 的子元素,映射 XML Schema 模式为:

```
<xsd:element name="DB">
  <complexType>
    <all>
      <xsd:element name="book" type="BOOKTYPE"/>
      <xsd:element name="customer" type="CUSTOMER-
        TYPE"/>
    </all>
```

```
</complexType>
</xsd:element>
```

(2) 关系数据库的结构与 XML 文档结构完全不一样,后者是基于层次结构或树型结构,现先将关系表结构转换如图 3 和图 4 所示。

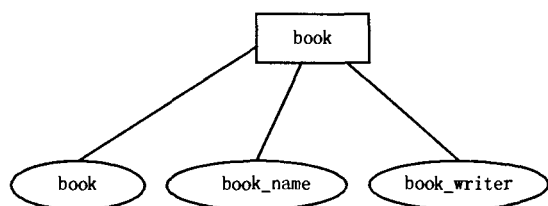


图 3 book 实体关系图

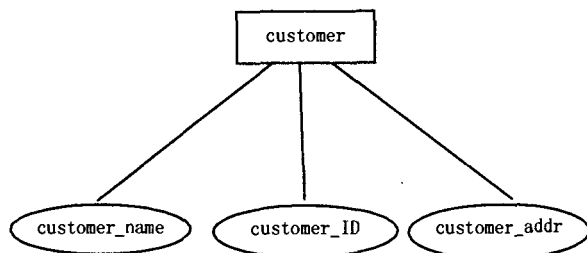


图 4 customer 实体关系图

将关系表名映射为 XML Schema 元素名,关系表中的属性转换为该元素的子元素或属性,现映射为:

```
<xsd:element name="book" type="BOOKTYPE">
<complexType>
<sequence>
<xsd:element name="book_ID" type="book_IDTYPE">
<xsd:element name="book_name" type="xsd:string">
<xsd:element name="book_writer" type="book_writer-
TYPE">
</sequence>
</complexType>
</xsd:element>
<xsd:element name="customer" type=
"customerTYPE">
<complexType>
<sequence>
<xsd:element name="customer_name" type="book_ID-
TYPE">
<xsd:element name="book_ID" type="book_IDTYPE">
<xsd:element name="customer_addr" type="xsd:token">
</sequence>
</complexType>
</xsd:element>
```

2.3 基于关系模式转换到 XML Schema 的约束

良好的关系数据库是满足 3NF(第三范式)和数据库相关约束,为了实现转换到 XML Schema 模式就要通过一系列映射规则包括主关键字约束、外关键字约束、值约束、域约束、非空约束和缺省值等。

(1) 主关键字约束和外关键字约束。

关系模式中表主键和外键的映射是关系模式参照完整性的保证。为了数据的完整性,book_ID 既是 book 关系表的主键,又是 customer 的外键,在转换为 XML Schema 模式时,使用“ID”和“IDREF”属性来实现。用“ID”属性映射主键如下:

```
<xsd:element name="book_ID">
<xsd:complexType>
<xsd:simpleContent>
<xsd:extension base="xsd:string">
<xsd:attribute name="id" type="xsd:ID"/>
</xsd:extension>
</xsd:simpleContent>
</xsd:complexType>
</xsd:element>
```

用“IDREF”属性映射外键如下:

```
<xsd:element name="book_ID">
<xsd:complexType>
<xsd:attribute name="idref" type="xsd:IDREF"/>
</xsd:complexType>
</xsd:element>
```

(2) 值约束和域约束。

关系数据库中的数据往往是约束的,例如 book_writer 是字符类型,长度不超过 4,在转换为 schema 模式可以这样:

```
<xsd:simpleType name="book_writerTYPE">
<xsd:restriction base="xsd:string">
<xsd:maxLength value="4"/>
</xsd:restriction>
</xsd:simpleType>
```

(3) 非空约束。

如果 book_name 是字符类型,可以为空,那么使用属性 nillable,其值“true”表示字符可以为空,值为“false”,则不能为空。例如:

```
<xsd:simpleType name="book_nameTYPE">
<xsd:restriction base="xsd:string">
<xsd:nillable="true"/>
</xsd:restriction>
</xsd:simpleType>
```

(4) 缺省值。

默认值和固定值通过给空元素增加值的方式来扩展实例^[5]。如果文档中存在空的元素,模式处理器根据模式文档的定义,会插入默认值或固定值。在 XSDL 中,默认值和固定值分别通过 default 和 fixed 属性设置。两个属性只能出现其中之一,它们是互斥的。如果元素为空,就填入默认值。下例中,声明了 book_

(下转第 208 页)

据库中的 ECA 规则来描述 workflow 中的任务处理,将事件作为唯一的过程集成与协调机制,基于事件触发规则执行,从而实现 workflow 任务的自动路由。完整描述一个事件需要一些参数,如事件的形式、名字、发生时间、传递的信息以及与其它事件的关系等^[5]。

文中引入了三种事件:任务状态相关事件、数据相关事件、时间相关事件。其中,任务状态相关事件由 workflow 中任务状态的变化而产生;数据相关事件由对 workflow 中数据流的操作而产生;而时间相关事件则是由 workflow 引擎对系统时间的监控产生^[6]。为了实现事件驱动的规则处理技术,还需要研究有效的事件驱动控制模型,使得系统中事件的发生可以很快被发现,从而触发执行相应动作。事件接口实时接收来自外界的消息,并转化成系统要求的格式存入事件库中,事件库中存储了信息家电发生的各种事件,当事件满足规则库中某条规则的条件,推理机便可得出结论,触发相应规则。workflow 执行过程中可能会产生新的事件,系统会将这些新的事件作为消息存入事件库中。图 3 是对事件驱动分布式 workflow 模型描述。

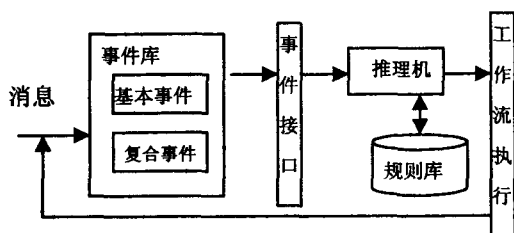


图 3 事件驱动分布式 workflow 控制模型

根据 workflow 管理的特点,本系统所采用的是正向推理,即从系统的首任务出发,引用规则库中的规则,向后寻找满足条件的任务或子过程,并执行产生能触

发后续任务的条件,递归调用规则,直至所有的任务或子过程被完成。

在专家系统的推理过程中,推理机每一循环中的可用规则往往不止一条,确定可用规则中可被启用的规则方法称为冲突消解策略。在冲突消解策略上,本系统采用优先级排序策略,谁的优先级高谁先执行,并优先获得资源。

4 结束语

文中针对传统 workflow 缺乏主动性、智能性等问题,借鉴了专家系统在知识表示、推理以及控制上成熟的理论,构建了基于专家系统的工作流系统模型,并结合事件驱动的分布式 workflow 技术,将之应用在智能家居的研究中,较好地解决了信息家电的异构性等问题,并实现了不同家电间的智能协作。

参考文献:

- [1] 范玉顺. 工作流管理技术基础[M]. 北京:清华大学出版社,2001.
- [2] Workflow Management Coalition. Workflow management coalition terminology and glossary[R]. Brussels: Workflow Management Coalition, 1996.
- [3] 张 鹏,吕汉新. 一种基于 CORBA/Java 的网络智能专家系统[J]. 计算机技术与发展, 2006, 16(12): 94-97.
- [4] 林尧瑞,张 钺,石纯一. 专家系统原理与实践[M]. 北京:清华大学出版社,1988.
- [5] 李 宁. 一种支持分布式 workflow 模型的工作流管理系统[J]. 计算机应用研究, 2005(3): 205-213.
- [6] 王新军. 事件驱动的分布式 workflow 技术在医疗保险信息系统中的应用[J]. 计算机工程与应用, 2005(23): 224-232.

(上接第 204 页)

writer 元素,并指定了默认值为“佚名”。

```
<xsd:elementname="book_writer" type="xsd:string" default="佚名" />
```

3 结束语

目前电子商务中的数据表达形式繁多,没有统一的格式,从而使不同企业之间信息交换成了障碍。XML Schema 具有清楚、灵活的描述 XML 信息内容和结构的能力,是电子商务开发应用的关键性技术。基于 XML Schema 技术特点,利用规范约束^[6]的 XML 实现了关系模式向 XML 模式的完全转化,而且能验证商务文档的有效性和为信息存取提供了保障机制;实现了系统具有跨平台、松耦合、异步性、可靠性、可扩展性等特点^[7]网上数据的交换、传输等应用。从而为关系数据在电子商务中数据转换提供更有效的支持。

参考文献:

- [1] W3C. XML. W3C Recommendation[EB/OL]. 2006-08. <http://www.w3.org/TR/REC-xml/>.
- [2] W3C. XMLSchema. W3CRecommendation[EB/OL]. 2004-10. <http://www.w3.org/XML/Schema.html>.
- [3] 马永恒,熊前兴. W3C XML Schema 模式的设计方法研究[J]. 计算机应用研究, 2006(5): 80-82.
- [4] Walmsley P. XML 模式权威教程[M]. 北京:清华大学出版社, 2003.
- [5] Graves M. Designing XML Databases[M]. 北京:机械工业出版社, 2002.
- [6] 何震瀛,李建中,王朝坤. 一种 XML 数据库的数据模型[J]. 软件学报, 2006, 17(4): 759-769.
- [7] 何 慧,陈 博. 基于 XML 和 JMS 的异构数据交换集成的研究[J]. 计算机技术与发展, 2006, 16(2): 70-72.