

基于 XML 的电子病历存储系统研究

吴万春 梁珂

(重庆医科大学附属儿童医院信息中心 重庆 400014)

〔摘要〕 电子病历应用是医院信息化的发展趋势, 基于 XML 的电子病历系统在国内已经逐步从研究走向了应用阶段。以 XML 作为电子病历的基本存储模型, 重点讨论该模型的设计以及实现方法, 在此基础上介绍 XML 作为电子病历存储基本单元的理论依据, 以及医院信息系统中关系数据模型与 XML 文档数据模型相互转换的实现手段。

〔关键词〕 XML; 电子病历; 关系数据库

Research on the Electronic Medical Records Storage System Based on XML WU Wan - chun, LIANG Ke, Information Center of Children's Hospital Affiliated to Chongqing Medical University, Chongqing 400014, China

〔Abstract〕 Electronic medical records (EMR) application is the development trend of hospital informatization. XML based EMR system is gradually running from the research to the application stage in China. The paper takes XML as the basic storage model, emphatically discusses its design and implementation method. On this basis, it introduces the theoretical basis of adopting XML as the basic storing unit for the electronic medical records, and explores implementation approaches of mutual transformation between relation data model and XML text data model in hospital information systems.

〔Keywords〕 XML; Electronic medical records; Relational database

再次在检索数据方面要快速、准确。

1 引言

在临床信息系统的应用中, 所有的临床信息最终都要反映到患者病历中。病历信息化在医院信息化建设中处于核心的地位, 电子病历是医院数字化之本^[1]。为满足临床和科研工作需要, 电子病历首先要保持病历的整体结构完整; 其次, 格式要灵活通用, 能对病历的结构和内容进行设置和修改;

2 XML 语言与电子病历

病历的描述模型是电子病历的最基础问题, XML 为病历内容的描述提供了有效手段。XML 是一种标识语言, 更是一种可以定义描述对象结构的元语言。XML 文档自含结构, 使得系统间交换的信息可以互相“理解”^[2,3]。使用 XML 作为病历内容的描述语言主要有以下优点: 可方便实现电子病历的结构化存储; XML 是一种可以定义描述对象结构的元语言^[4], 采用 XML 技术解决电子病历结构化问题给电子病历系统的应用带来大量好处^[2]; 方便电子病历文档的生成与修改; 在获取电子病历基础

〔收稿日期〕 2009 - 12 - 04

〔作者简介〕 吴万春, 硕士, 副主任, 主要研究方向为计算机系统结构、办公自动化系统等, 发表论文数篇。

数据之后, 利用生成模块生成 XML 文档, 再利用编辑工具直接进行病历的修改与文字录入, 完成后可以送入复审流程, 也可以直接归档^[5]; 有利于通用查询与统计模块的实现: 电子病历的结构化和知识表示的标准化, 使得针对电子病历的统计或查询简单易行, 也使得医疗宏观管理工作变得容易。

3 基于 XML 的电子病历存储系统设计

3.1 电子病历文书存储设计

3.1.1 利用 XML 存储电子病历的依据 XML 的建模方法是一种高度灵活的、能够表达同类而异构数据的模型^[4]。另外, XML 表达具有良好的扩展性^[6]。XML 可以在不破坏现有结构和系统的情况下增加新的数据字段。若改变数据模型, 只需改变数据模式定义, 不需要重新编写现有的程序。

3.1.2 利用 XML 存储电子病历的原理 要先定义病历内容的结构 (DTD)。在此基础上实现业务数据库中病人信息到病历结构的转换, 从而实现由以类型为中心的数据库描述到以病人为中心的描述, 形成的 XML 文档是病历存储管理的基本单位^[7]。采用 XML 文档来记录病历并不排斥病人信息的数据库表示。XML 描述与数据库记录有各自的适用范围, 可共同存在并保持相对独立。数据库系统主要用于支持日常的业务处理和病人信息的采集, 需要大量的数据检索和更新。业务系统要保持很高效率和稳定性, 大量纯文本的描述信息无法很好地存储在关系数据库中, 也不利于关系数据库的检索和更新, 对关系数据库系统的效率和数据存储方式都会带来影响。基于 XML 文档的电子病历系统所实现的病历存储, 主要是对病人信息的提取, 对病史、病程的客观文字性描述, 对一些不适于建模为关系数据库的数据结构进行存储。

3.2 关系数据库到 XML 文档的转换

从关系数据库到 XML 文档转换的主要目的是保存病人病历的历史记录, 减少不必要的重复录入工作, 同时可以为实现区域电子病历共享提供基本数据接口和原始数据。由于各医院的医疗信息系统

的医疗数据多数采用关系数据库的方式存储, 为了方便共享, 为区域医疗服务, 需要实现关系数据库向 XML 文档的转换。基本转换步骤, 见图 1。

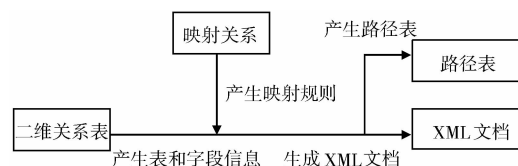


图 1 关系数据库转换为 XML 文档

3.2.1 对关系数据库中的表进行标准化命名 目的是实现不同表中的意义相同的字段的名称统一, 从而确保字段名在整个关系集合中的一致性。这样可以避免相同字段信息的不同元素映射, 减少重复映射。

3.2.2 选择映射的范围 由于关系数据库中的表很多, 涉及的信息包括各个方面, 所以如果将整个数据库中的表映射为一个 XML 文档, 则此文档的信息会十分混乱。因此可以选定 XML 文档数据要描述的主题, 确定该主题所包括的关系表, 将这些关系表转换成一个 XML 文档, 这样就会显得十分清晰。

3.2.3 获取关系表的所有信息 确定了要映射的关系表的范围后, 应该得到每个表的相关信息, 如包括的字段名称、字段的类型、字段的长度、字段是否为空、字段的缺省值、主键的构成、外键的构成、与此表发生联系的表的名称、各个索引的构成、各个表之间的一对一或一对多关系。这些信息往往可以通过关系数据库中的系统表得到。

3.2.4 确定根元素和基本元素 XML 文档的根元素, 可以命名为主题的名字, 先将所有的表都映射为一个对应的元素, 元素名可以直接使用表名。这些元素均作为根元素的子元素。将关系表分为两种: 一种是基本内容表, 一种是辅助表 (代码表及中间表等等)。这两种表的功能不同, 映射的对象也有所不同。

3.2.5 字段的映射 每个表的字段既可以映射为子元素, 又可以映射为属性。主要的原则是构成主键和外键及索引的字段, 往往反映一些基本性质且

变化较小,所以应映射为属性。而其它的字段为了保证将来的可扩充性,可以映射为子元素。

3.2.6 辅助表的处理 将其映射为一个元素后,可以作为引用它的表所对应的元素的子元素,且这个辅助表映射成的元素应为枚举类型,将辅助表中的数值列在枚举值集合。关系型记录可以被格式化为元素,其数据作为属性值或作为嵌入元素。把记录格式化为属性值是比较简单的,但把记录格式化为元素更通用一些^[4]。

3.2.7 生成 XML 文档 根据以上映射转换规则,生成 XML 文档,并且把 XML 文档的存放路径、创建者、修改者等相关信息存入路径表。

3.3 XML 文档到关系数据库的转换

从 XML 文档到关系数据库转换的主要目的是保存与疾病特征、病人基本信息等相关的关键性词语。把这类重要的关键性词语保存在二维的关系数据库中,可以提高查询统计的效率,为科研教学提供大量的原始基础数据。系统需要将这些集成了病历信息的 XML 文档进行解析,根据自定义映射文件的规则进行映射,存入关系数据库中。由于 XML 文档在结构上与关系数据库表格之间有很大差异,所以映射文件主要是定义了这两种结构之间的映射规则,用于实现数据从 XML 文档到数据库表格的转移。基本转换步骤,见图 2。

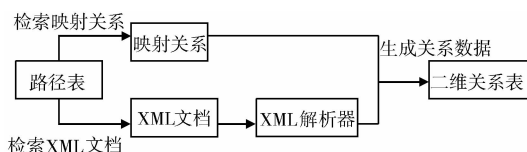


图 2 XML 文档转换为关系数据库

3.3.1 获取映射文件和 XML 文档 读取数据库中路径表的内容,获取 XML 文档和映射文件的存放路径。遍历其路径下的文件夹,读取 XML 文档和相应的映射文件。

3.3.2 解析映射文件和 XML 文档 解析 XML 文档,生成 DOM 树。解析映射表,得到从 XML 文档到数据库表格结构的映射规则。利用第 3 方的 XML

解析器工具所提供的一系列解析包解析 XML 文档,得到所有的元素和属性等信息,对映射表进行解析,得到相应的映射规则。

3.3.3 根据映射规则存储数据 把解析 XML 文档得到的数据和解析映射文件得到的映射规则进行对应,调用 SQL 语句把解析得到的 XML 文档的数据保存于关系数据库的表格中,而存储 XML 的目的就是为了查询统计。把病人的病历信息存入数据库后,可以用 SQL 语句取出数据,实现查询或转换为 XML 形式与其它数据库交换数据。

4 结语

电子病历系统是实现医务人员业务流程全面信息化的临床信息系统,也是提高医疗质量,实行规范化、智能化、标准化医院管理的重要工具之一。本文着重讨论了基于 XML 电子病历系统的存储体系,这为医疗机构之间共享数据提供了一种基本的技术手段,对发展区域化医疗具有重要意义。

参考文献

- 1 范启勇,杰宁,邬惊雷. 数字化医院的内涵特征和基本框架 [J]. 中华医院管理杂志, 2006, 22 (1): 57-58.
- 2 Hongyan Li, Ming Xue, Ying Ying. A Web-based and Integrated Hospital Information System [J]. The Digital Hospital, 2004, (4): 30-36.
- 3 聂培尧,李战怀,胡正国. 基于 OEM 的 XML 半结构数据的模式描述方法 [J]. 计算机工程与设计, 2003, 24 (1): 9-12.
- 4 许斗,陈恩红. XML 的半结构化数据表示方法及其在医学文档处理中的应用 [J]. 计算机工程, 2002, 8 (1): 57-58.
- 5 陈嗣棋,吕旭东,段会龙. 基于 XML 技术的医疗电子病历应用体系的研究 [J]. 中国医疗器械杂志, 2004, 28 (4): 21-22.
- 6 张兆国,薛万国,包鹏飞,等. 建设基于 XML 技术电子病历系统的核心价值 [J]. 中国数字医学, 2009, 2 (3): 11.
- 7 宋善德,肖必强. 关系模式下的 XML 数据存取技术研究 [J]. 计算机工程与科学, 2004, (8): 63.