一种XML数据库的数据模型

何震瀛, 李建中, 王朝坤

何震瀛, 李建中, 王朝坤

(哈尔滨工业大学 计算机科学与技术学院,黑龙江 哈尔滨 150001)

作者简介: 何震瀛(1977－),男,黑龙江牡丹江人,博士生,主要研究领域为XML数据管理.李建中(1950－),男,博士,教授,博士生导师,CCF高级会员,主要研究领域为数据库技术,并行计算技术.王朝坤(1976－),男,博士生,主要研究领域为音频数据管理,P2P数据管理.

联系人: 何震瀛 Phn: +86-451-86415872, E-mail: hzy@hit.edu.cn, http://db.cs.hit.edu.cn

Abstract
Data model is a key research field in the community of XML data management. However, the existed works fall short in their ability to model complex data structure of XML database or to support complete operations. This paper proposes a new mapping-based data model. This model aims to give the precisely defined notations for complex data structure and semantics of XML database. Along with the model, this paper also presents an associated algebra formally which includes a set of path expression operations and data modification operations. This model has already adopted in an XML-based information integration system. It shows that this model can give the precise semantics of XML database, and support the XML database applications effectively.

摘要
数据模型是XML数据管理领域研究的核心问题之一.现有的数据模型在表达XML数据库复杂的数据结构和操作方面仍有不足.以映射为基础,提出了一种新的数据模型.该数据模型给出了XML数据库复杂的数据结构和语义的精确定义,并提供了数据结构上操作代数的定义,包括路径表达式操作和数据维护操作.该数据模型已应用于一个基于XML的信息集成系统中.事实表明,它能够有效地支持XML数据管理的应用.

References:

[1] Quass D, Widom J, Goldman R, Haas K, Luo Q, McHugh J, Nestorov S, Rajaraman A, Rivero H, Abiteboul S, Ullman J, Wiener J. LORE: A lightwight object repository for semistructured data. In: Jagadish HV, Mumick IS, eds. Proc. of the 1996 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 1996. 549.

[2] Zhou AY, Lu HJ, Zheng SH, Liang YQ, Zhang L, Ji W, Tian ZP. VXMLR: A visual XML-relational database system. In: Apers P, Atzeni P, Ceri S, Paraboschi S, Ramaohanarao K, Snodgrass R, eds. Proc. of the 27th Int'l Conf. on Very Large Data Bases. San Fransisco: Morgan Kaufmann Publishers, 2001. 719-720.

[3] Wang Q, Zhou JM, Wu HW, Xiao JC, Zhou AY. Mapping XML documents to relations in the presence of functional dependencies. Journal of Software, 2003,14(7):1275-1281 (in English with Chinese abstract). http://www.jos.org.cn/1000-9825/14/1275.htm

[4] Florescu D, Kossmann D. Storing and querying XML data using an RDBMS. IEEE Engineering Bulletin, 1999,22(3):27-34.

[5] Tatarinov I, Viglas S, Beyer K, Shanmugasundaram J, Shekita E, Zhang C. Storing and querying ordered XML using a relational database system. In: Franklin M, Moon B, Ailamaki A, eds. Proc. of the 2002 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 2002. 204-215.

[6] Yoshikawa M, Amagasa T, Shimura T, Uemura S. XRel: A path-based approach to storage and retrieval of XML documents using relational databases. ACM Trans. on Internet Technology, 2001,1(1):110-141.

[7] Zhang C, Naughton J, DeWitt DJ, Luo Q, Lohman G. On supporting containment queries in relational database management systems. In: Aref W, ed. Proc. of the 2001 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 2001. 426-437.

[8] Arenas M, Libkin L. A normal form for XML documents. ACM Trans. on Database Systems, 2004,29(1):195-232.

[9] Christophides V, Cluet S, Moerkotte G, Siméon J. On wrapping query languages and efficient XML integration. In: Chen W, Naughton J, Bernstein P, eds. Proc. of the 2000 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 2000. 141-152.

[10] Ludascher B, Papakonstantinou Y, Velikhov P. Navigation-Driven evaluation of virtual mediated views. In: Zaniolo C, Lockemann P, Scholl M, Grust T, eds. Advances in Database Technology-EDBT 2000, 7th Int'l Conf. on Extending Database Technology. Berlin, Heidelberg: Springer-Verlag, 2000. 150-165.

[11] Fankhauser P, et al. XQuery 1.0 and XPath 2.0 formal semantics. 2005. http://www.w3.org/TR/query-semantics/

[12] Chamberlin D, et al. XQuery 1.0: An XML query language. 2005. http://www.w3.org/TR/xquery/

[13] Buneman P, Fernandez M, Suciu D. UnQL: A query language and algebra for semistructured data based on structural recursion. The VLDB Journal, 2000,9(1):76-110.

[14] Beeri C, Tazban Y. SAL: An algebra for semistructured data and XML. In: Cluet S, Milo T, eds. Proc. of the 2nd ACM SIGMOD Workshop on the Web and Databases. INRIA, 1999. 37-42. http://www-rocq.inria.fr/~cluet/webdb99.html

[15] Jagadish H, Lakshmanan L, Srivastava D, Thompson K. TAX: A tree algebra for XML. In: Ghelli G, Grahne G, eds. Proc. of the 8th Int'l Workshop on Database Programming Languages. Berlin, Heidelberg: Springer-Verlag, 2001. 149-164.

[16] Fernandez M, Siméon J, Wadler P. An algebra for XML query. In: Kapoor S, Prasad S, eds. Proc. of the Foundations of Software Technology and Theoretical Computer Science, 20th Conf. Berlin, Heidelberg: Springer-Verlag, 2001. 11-45.

[17] Meng XF, Luo DF, Jiang Y, Wang Y. OrientXA: An effective XQuery algebra. Journal of Software, 2004,15(11):1648-1660 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/15/1648.htm

附中文参考文献:
[17]　　孟小峰,罗道锋,蒋瑜,王宇.OreintXA:一种有效的XQuery查询代数.软件学报,2004,15(11):1648-1660.　　http://www.jos.org.cn/1000-9825/15/1648.htm