半结构化查询重写的MiniCon算法

陶 春, 汪 卫, 施伯乐

陶 春, 汪 卫, 施伯乐
(复旦大学 计算机与信息技术系,上海 200433)
作者简介: 陶春(1974－),男,湖北麻城人,博士生,主要研究领域为数据库,知识库;汪卫(1970－),男,博士,教授,主要研究领域为数据库技术,数据挖掘,XML数据管理;施伯乐(1935－),男,教授,博士生导师,主要研究领域为数据库系统及其应用,数据仓库与数据挖掘,数字图书馆,安全数据库.
联系人: 陶 春 Phn: +86-21-65642219, Fax: +86-21-65642219, E-mail: ctao@fudan.edu.cn, http://www.fudan.edu.cn

Abstract

This paper addresses the semi-structured query rewriting problem for TSL (tree specification language), a language for querying semi-structured data. An algorithm that can find the maximally-contained rewriting query is presented, when a semi-structured query and a set of semi-structured views are given. The idea is borrowed from MiniCon, a scalable relational query rewriting algorithm, and some new problems for semi-structured query rewriting, e.g., object-id dependency and set value variable mapping, are solved. It is shown that the algorithm is correct.

Tao C, Wang W, Shi BL. MiniCon algorithm for semi-structured query rewriting. *Journal of Software*, 2004,15(11):1641~1647.
http://www.jos.org.cn/1000-9825/15/1641.htm

摘要

研究了基于半结构化数据查询语言TSL(tree specification language)的查询重写问题.提出了一种半结构化查询重写算法,解决了在给定一个半结构化查询和一组半结构化视图的情况下,找到最大被包含重写的问题.算法借用了可伸缩的关系查询重写的MiniCon算法的思想,解决了半结构化数据模型之下查询重写的一些新问题(如标识符依赖、集合值变量映射等).证明了算法的正确性.

References:

[1] Halevy AY. Answering queries using views: a survey. The VLDB Journal?The Int'l Journal on Very Large Data Bases, 2001, 10 (4):270~294.

[2] Pottinger R, Halevy AY. MiniCon: A scalable algorithm for answering queries using views. The VLDB Journal?The Int'l Journal on Very Large Data Bases, 2001,10(2-3):182~198.

[3] Abiteboul S, Buneman P, Suciu D. Data on the Web: From Relations to Semistructured Data and XML. San Francisco: Morgan Kaufmann Publishers, 2000.

[4] Papakonstantinou Y, Vassalos V. Query rewriting for semistructured data. In: Proc. of the 1999 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 1999. 455~466.

[5] Gao J, Tang SW, Yang DQ, Wang TJ. Query rewriting for semi-structured data. Journal of Computer Research and Development, 2002,39(2):165~171 (in Chinese with English abstract).

附中文参考文献：
[5] 高军,唐世渭,杨冬青,王腾蛟.半结构化数据查询重写.计算机研究与发展,2002,39(2):165~171.

附中文参考文献：
[5] 高军,唐世渭,杨冬青,王腾蛟.半结构化数据查询重写.计算机研究与发展,2002,39(2):165~171.