Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Visual Languages and Computing
17 (2006) 25–45
Journal of Visual Languages & Computing
www.elsevier.com/locate/jvlc

VIREX: visual relational to XML conversion tool

Anthony Lo[a], Reda Alhajj[a,b,*], Ken Barker[a]

[a]*ADSA Lab, Department of Computer Science, University of Calgary, Calgary, Alta., Canada*
[b]*Department of Computer Science, Global University, Beirut, Lebanon*

Received 2 May 2004; received in revised form 7 August 2005; accepted 15 August 2005

Abstract

Developing user-friendly transformation tools for converting all or part of a given relational database into XML has not received enough consideration. This paper presents a flexible user interface called VIREX (VIsual RElational to XML), which facilitates converting a selected portion of a given underlying relational database into XML. VIREX works even when the catalogue of the underlying relational database is missing. For the latter case, VIREX extracts the required catalogue information by analyzing the underlying database content. From the catalogue information, whether available or extracted, VIREX derives and displays on the screen a graph similar to the entity-relationship diagram. VIREX provides a user-friendly interface to specify on the graph certain factors to be considered while converting relational data into XML. Such factors include: (1) selecting the relations/attributes to be converted into XML; (2) specifying a predicate to be satisfied by the information to be converted into XML; (3) deciding on the order of nesting between the relations to be converted into XML. All of these are specified by a sequence of mouse clicks with minimum keyboard input. As a result, VIREX displays on the screen the XML schema that satisfies the specified characteristics and generates the XML document from the underlying relational database. Finally, VIREX is essential to optimize the amount of information to be transferred over a network by giving the user the flexibility to specify the amount of relational data to be converted into XML. Also, VIREX can be used to teach XML to beginners.

Keywords: Data conversion; Flexible interface; Graphical user interface; Relational database; Visual query language; XML

*Corresponding author. ADSA Lab, Department of Computer Science, University of Calgary, Calgary, Alta., Canada. Tel.: +1 403 2109453.

E-mail addresses: chiul@cpsc.ucalgary.ca (A. Lo), alhajj@cpsc.ucalgary.ca (R. Alhajj), barker@cpsc.ucalgary.ca (K. Barker).

1045-926X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jvlc.2005.08.002

1. Introduction

XML (eXtensible Markup Language) has emerged as a recommendation of W3C and is being gradually accepted as the standard format for data exchange between different platforms and partners [1]. Among its great advantages are portability, extensibility and the possibility to add semantics, in particular structural constraints, to data within the document itself. The problem of converting relational data into XML assumes special significance because huge amount of data is currently stored in relational databases and relational databases are still dominant. Some accomplishments have already been done in this direction and there are several tools and languages described in the literature to generate XML document from relational database, e.g., [2,3]. However, most of these approaches require users to learn a new language before they can query the database to create the desired data file. Such approaches mainly focus on finding XML schema (e.g., (DTD) [4] and XML schema [5]) that best describes a given relational database with a corresponding well-defined database catalogue that contains all the necessary information about tables, attributes, keys, constraints, etc. DTD uses a regular expression like language to specify the structure of the document. Disadvantages of using DTD include its limited data types and its very limited ability to add constraints. On the other hand, XML schema allows for a more sophisticated and precise definition of both the structure and the data.

In general, relational-to-XML conversion is not trivial because the two data models are significantly different. While relational data are flat, normalized and the schema is often proprietary, XML data could be nested, unnormalized, and its schema is public. The relational-to-XML conversion involves mapping relational tables and attributes into XML elements and attributes, creating XML hierarchies, and processing values in an application-specific manner. Such conversion process is systematic and mainly performed by professionals. However, we argue that it is essential to consider visual user-friendly interface to help in the conversion process. We have been motivated by the fact that providing such an interface will definitely increase the acceptance of XML in the community and will provide an excellent training tool for XML beginners.

During the past two decades, extensive research efforts have successfully produced different visual user interfaces for querying relational and object-oriented databases; they are well received by naive users. However, developing user-friendly transformation tools that provide the opportunity to convert all or part of a given relational database into XML has not received enough consideration yet.

This paper presents a flexible user interface called VIREX (VIsual RElational to XML), which facilitates converting all or part of a given relational database into XML. VIREX connects to the relational database specified by the user; it then uses the information from the database catalogue (if available) to derive and display on the screen a graph similar to the entity-relationship diagram. VIREX works even when the catalogue of the underlying relational database is missing. For the latter case, VIREX extracts the required catalogue information by employing our approach described in [6] for analyzing the underlying database content.

The graph simply summarizes all schema information in a visual interface that shows the relations, their attributes and the links between them. VIREX provides a user-friendly interface to specify on the graph certain factors to be considered while converting relational data into XML. Such factors include: (1) selecting the relations/attributes to be converted into XML; (2) specifying a filtering predicate to be satisfied by the information