Chapter VIII

An Algorithm for Transforming XML Documents Schema into Relational Database Schema

Abad Shah, University of Engineering & Technology (UET), Pakistan
Jacob Adeniyi, King Saud University, Saudi Arabia
Tariq Al Tuwairqi, King Saud University, Saudi Arabia

Abstract

The Web and XML have influenced all walks of lives of those who transact business over the Internet. People like to do their transactions from their homes to save time and money. For example, customers like to pay their utility bills and other banking transactions from their homes through the Internet. Most companies, including banks, maintain their records using relational database technology. But the traditional relational database technology is unable to provide all these new facilities to the customers. To make the traditional relational database technology cope with the Web and XML technologies, we need a transformation between the XML technology...
and the relational database technology as middleware. In this chapter, we present a new and simpler algorithm for this purpose. This algorithm transforms a schema of a XML document into a relational database schema, taking into consideration the requirement of relational database technology.

Introduction

An electronic document on the Web contains regular and irregular structures that may not be understood by users (Suciu, 1999; Abiteboul & Vianu, 1997; Brayan, 1997). Such a document (or data) is referred to as semistructured data (Suciu, 1999; Abiteboul, 1997). Contrary to the data in relational databases (RDBs), the semistructured data is stored without any schema or with a vague schema (Buneman, 1997; Suciu, 1999; Buneman, 1997). Beside the Web, there are many other sources of semistructured data, such as heterogeneous networking of integrated systems, file systems, electronic mail systems, digital libraries, etc. (Abiteboul, 1997; Buneman, 1997).

The introduction of Extensible Markup Language (XML) as a standard data/information representation has facilitated the publication of electronic data on the Web (W3C, 2003). This language also provides a hierarchical format for the data exchange over the Web with structure (St. Laurent, 1999; Bray, Paoli, Sperberg-McQueen, & Maler, 2002). Information in an XML document is represented as nested element structures, which start with a root element. An element can have an attribute or a sub-element (for further details about XML see W3C (2003) and Bray et al. (2002)). An XML document has an optional part, which is called Document Type Declaration/Description (DTD). A DTD of an XML document is considered as the schema of the XML document (W3C, 2003; Bray et al., 2002; Men-Hin & Fu, 2001).

A relational database (RDB) has two main components, a schema and data files (or operational files) which are created according to the schema. As said earlier, a DTD is considered as a schema of an XML document, but there are noticeable differences between a RDB schema and an XML document schema (DTD). We give a complete comparison between them in Table 1. The basic difference between them is that a DTD represents a hierarchical structure whereas a RDB schema represents a relational (tabular) structure. We can consider XML documents schema analogous to the classical hierarchical data model.

XML is considered as the best tool for representing and exchanging information on the Web (St. Laurent, 1999; Bray, Paoli, Sperberg-McQueen & Maler, 2002). The language allows users to define and also display data on the Web. These features make XML powerful and different from Hypertext Markup Language.
Related Content

Epistemetrics: Why We Measure
www.igi-global.com/chapter/epistemetrics-measure/24878?camid=4v1a

Analysis of Click Stream Patterns using Soft Biclustering Approaches
www.igi-global.com/article/analysis-click-stream-patterns-using/51368?camid=4v1a

Information Problems in Competitive Markets and Their Impact on Labor Markets
Hasan Bilgehan Yavuz (2019). Global Information Diffusion and Management in Contemporary Society (pp. 189-211).
www.igi-global.com/chapter/information-problems-in-competitive-markets-and-their-impact-on-labor-markets/208072?camid=4v1a
Challenges and Opportunities for Innovation in Teaching and Learning in an Interdisciplinary Environment
www.igi-global.com/chapter/challenges-opportunities-innovation-teaching-learning/68334?camid=4v1a