

Foundations of a New Proposal for Querying Relational Databases and XML Documents

Ana Fermoso García¹, María José Gil Larrea², Luis Joyanes Aguilar¹, and Jesús Luis Díaz Labrador²

¹Facultad de Informática. Universidad Pontificia de Salamanca, Compañía 2, 37002 Salamanca, Spain

afermoso@upsa.es, ljoyanes@fpablovi.org
<http://www.upsa.es>

²Facultad de Ingeniería. Universidad de Deusto, Avda. Universidades 24, 48007 Bilbao, Spain

{marijose, josuka}@eside.deusto.es
<http://www.deusto.es>

Abstract. Today, a large amount of information has been stored in XML, which has become the most useful standard of data interchange in the web and e-business world. Nonetheless, a large quantity of information remains stored in Relational Databases.

It is very important to be able to jointly manage these two data formats. One solution would be an integrated query model that would at the same time allow the query of a relational database and of XML data. Because of the growing number of applications that have to work in a web environment, the proposed system should use a query language based in XML and also return answers in the same XML format.

After having made a detailed study, in this paper we make comparing different systems that link both, databases and XML. Besides, we propose a new system to overcome the limitations found in them.

1 Introduction

This paper tries to center in three aspects. First, a specific kind of database, Relational databases, because it is one of the most used methods to store and manage data. Second, the relation between the XML world, because XML has reached the status of standard of exchange and data representation in the web, and the relational database one. Finally, to query the two kinds of information at the same time with the same system and in an efficient way. That is to say, in the end the user makes the query without knowing what the real data source queried is.

2 A New Database Query System Based on XML

The main aim is to create a system that allows querying relational database and XML documents using the same query system in an efficient way. Besides, this system should be XML based, in this way all the environment has the same appearance.

The first step would be to establish the model that allows to rewrite the relational database structure in a special format XML-accessible, but without any kind of translation of the data or database records, which was one of the problems of other

similar systems. The idea would be to translate the database data model, the relation-entity database model, to a special XML format, but without any change in the database data.

The second important step would be to establish the query XML based language, to access databases in this environment. Then we would define the model to translate the query from this language into SQL, to make the final access to the database using the Management Database System, which will be more efficient.

The query language should have enough power to extract information from XML documents. As XQuery [1] is not yet standardized nor implemented, we assume as a premise that XSL (Extensible Stylesheet Language), could serve us to get the intended goal. This language is defined [2], widely implemented, and nowadays an industry de facto standard.

To translate a database model into an XML model, we would have to map database elements such as tables, relations and fields, into concepts of XML as elements and attributes. The resulting XML document would appear as a XML Schema document.

In the second step we would have to translate the sentences of the query XML language selected, to the SQL language clauses.

Once we have the query in SQL, we can execute it over the database using the Database Management System. Finally, the returned results of the query are translated into XML.

The architecture of the proposed system would be made up of two independent components. The first one translates the data model in the database to a special XML Schema. The second component receives a query in a XML language over the database, which would have already been adapted to the system with the previous component, and gives out the query results in XML. Figure 1 shows this architecture format.

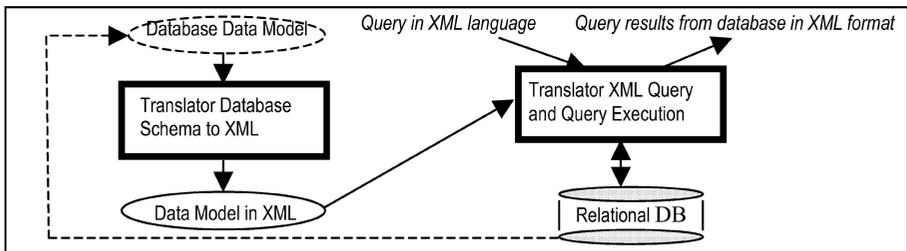


Fig. 1. Architecture of the XML Query System to Relational Databases

In this moment this system is in its final development process and have already been made tests that show it's suitable.

References

1. Scott et all. XQuery 1.0: An XML Query Language. <http://www.w3.org/TR/xquery/> (2002)
2. W3C World Wide Web Consortium The Extensible Stylesheet Language (XSL) <http://www.w3.org/Style/XSL/> (2003)