Database Systems CS6530 (Fall '08) Project Proposal

Pramod R Sanaga
School of Computing,
pramod@cs.utah.edu,
u0504960

September 22, 2008

The broad goal of this research project is to add an XQuery to SQL translation module to the existing codebase of Shrex [1].

Details:

Shrex is a tool to map annotated XML schema into equivalent relational schema, with the help of user input during the mapping process. It then allows XML documents to be shredded into an RDBMS and stored in terms of tables with columns as the attributes of the corresponding XML structure. Shrex implements a query translation module that takes XPath queries(only a subset of the language spec is currently implemented), turns them into SQL queries on the stored tables and evaluates them.

XQuery and XPath are in many respects similar in their underlying data model. XPath is a subset of XQuery, which is more expressive and allows for complex queries other than path based predicates. As far as this project is concerned, the FLWOR construct of XQuery is of the most interest. We would also look at implementing as many of the built-in functions of XQuery as possible.

Approximate timeline:

October 19th: Would finish reading the related work publications and the shrex codebase. Would add the ability to translate and evaluate simple path based predicates written in the XQuery language.

November 19th: Adding the ability to translate FLWOR constructs, including nesting if feasible.

December 19th: Looking at adding other constructs such as If-else, in-built functions and non-path based predicates.

Tools:

Shrex is written using Java. So, I will be using Java for the XQuery to SQL module. Depending on the extent of the existing query translation framework in Shrex, I might need to ues JavaCC to write a simple parser to process XQuery.

References

[1] S. Amer-Yahia, F. Du, and J. Freire. A comprehensive solution to the xml-to-relational mapping problem. In *WIDM '04: Proceedings of the 6th annual ACM international workshop on Web information and data management*, pages 31–38, New York, NY, USA, 2004. ACM.

[2] D. DeHaan, D. Toman, M. P. Consens, and M. T. Özsu. A comprehensive xquery to sql translation using dynamic interval encoding. In *SIGMOD '03: Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, pages 623–634, New York, NY, USA, 2003. ACM.

[3] T. Grust, M. Mayr, J. Rittinger, S. Sakr, and J. Teubner. A sql: 1999 code generator for the pathfinder xquery compiler. In *SIGMOD '07: Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, pages 1162–1164, New York, NY, USA, 2007. ACM.

[4] R. Krishnamurthy, R. Kaushik, and J. F. Naughton. XML-to-SQL query translation literature: The state of the art and open problems. In *Database and XML Technologies, First International XML Database Symposium (XSym)*, pages 1–18, 2003.