SHREXQUERY: AN IMPROVED QUERY TRANSLATOR FOR XPATH

Qing Zhang, Julianna Freire
School of Computing

The use of relational database management systems (RDBMSs) to store and query XML data has attracted considerable interest with a view to leveraging their powerful and reliable data management services. Due to the mismatch between the relational and XML data models, it is necessary to first shred and load the XML data into relational tables, and then translate XML queries over the original data into equivalent SQL queries over the mapped tables.

ShreX is an XML-to-relational mapping framework and system that provides the first comprehensive and end-to-end solution to the relational storage of XML data. Mappings in ShreX are defined through annotations to an XML Schema. The use of XML Schema simplifies the mapping process, since it does not require users to master a new specialized mapping language. The use of annotations allows mapping choices to be combined in many different ways. As a result, ShreX not only supports all the mapping strategies proposed in the literature, but also new useful strategies that had not been considered previously. The availability of a mapping specification makes it possible to develop generic components for both shredding documents and translating queries. The first public release of ShreX included a simple query translator, which only handles a limited subset of XPath (simple path expressions).

In this project, we will implement a new and improved translator, ShreXQuery. ShreXQuery will not only allow ShreX to handle a wider class of queries (i.e., XPath queries), but it will also attempt to generate efficient SQL queries. ShreXQuery will be included in the ShreX release (ShreX is freely available and can be downloaded at http://www.cse.ogi.edu/ShreX).