



UMass Lowell Computer Science Colloquium Announcement

Speaker: Prof. Murali Mani, WPI
Date & Time: Wednesday, March 21, 2007, 3:00pm--4:00pm
Place: Olsen 311, Refreshments are served at 2:40pm

Efficient XML Stream Processing: The Raindrop Approach

As streaming data becomes more and more prevalent, it is imperative that we develop techniques to process such data efficiently. In this talk, we will focus on XML stream processing and our system that we call Raindrop, which processes XQuery expressions over XML streams. XQuery can be considered as consisting of three parts: (a) pattern retrieval via XPath expressions (b) filtering via predicates and (c) output restructuring. There are at least two different opportunities for performing pattern retrieval -- (a) retrieve the patterns on the streaming XML tokens using automaton, (b) extract several tokens into an appropriate object, say DOM object, and perform navigation on this object. Which approach to use for retrieving the different patterns will depend on the query/XML stream characteristics. In Raindrop, we model pattern retrieval using automaton and DOM-based navigation uniformly using an algebraic paradigm, and use cost-based approaches for choosing an efficient plan. Further, using automaton for query processing is becoming more and more prevalent, and we shall examine trade-offs between different automata-based approaches for the same.

Acknowledgements: Raindrop is partially sponsored by NSF grant, IIS 0414567. Several people have been involved in this work: Elke Rundensteiner, Hong Su, Jinhui Jian, Ming Li, Mingzhu Wei, Andrei Vasilescu, Shoushen Wang, Drew Ditto, Bogomil Tselkov, Nikhil Sreenath, and others.

Bio:

Murali Mani finished his PhD in Computer Science from UCLA in 2003. Before this, he did his BTech in Computer Science and Engineering from IIT Madras in 1998, and his MS in Computer Science from UCLA in 2000. His prime research interests are in databases. He has worked extensively on XML technologies, in particular XML schema languages, and XML query processing.