

COMPSCI732 S1 C 2005 Assignment 1 XML tools and techniques

Worth: 16.6% towards your final grade
Due date: Fri 18th March at 1pm

Aim

Semistructured data, in particular XML, has emerged as one of the primary means for information exchange and content management [2]. In order to manage the data efficiently various tools and techniques are being researched and developed [1,3]. These include numbering schemes for answering queries, XML databases based on relational databases, access control mechanisms to ensure the data is secure, efficient querying mechanisms, materialized view mechanisms, and an understanding of keys in this new setting. Much of the inspiration for this work comes from relational databases. The aim of this assignment is to give you a deeper understanding of one particular area in this setting.

Task

In class you will be given a paper that was presented at the Australasian Database Conference (ADC). Your task is to review this paper, and write a report and powerpoint presentation of your review. The report should be at least 6-10 pages in length and the presentation should be 20 minutes in length. The target audience for both the report and the presentation is other graduate students.

There are a number of ways to find background and related work. We list only some of them here.

- Textbooks [4,5].
- The citations and references in the paper.
- A Google search with appropriate keywords.
- Digital libraries and bibliographies (DBLP).
- Related work carried out by the same author or same research group.
- CiteSeer.
- Google Scholar (<http://scholar.google.com/>).

Your report and presentation should be structured as follows:

- Title and authors of paper.
- Reviewed by with your name, UPI and student id.
- Abstract.
- Introduction.
- Background.
- Body.
- Conclusions.
- References.

Your report and presentation must provide sufficient background information for the target audience to understand your review. For example, if you are reviewing a paper about access control in XML Databases, you must explain what "access control" is. Your report and presentation should answer the following questions:

- What is the motivation for the work?
- What is the state-of-the-art in the research area being addressed?
- What is the research question the paper is addressing?
- What are the key findings of the paper?

All sections should include relevant references. Remember to cite work of others if you are using it in your report.

If you have any further questions about this assignment, please don't hesitate to either email me or come and see me.

Due dates and marking details

The review is due on Fri 18th March at 1pm. Both the report and presentation are to be submitted as pdf files. This assignment will count 16.6% towards your final grade.

The project will be marked on the clarity of your report and presentation, whether you have addressed the questions given above, whether you have demonstrated a good understanding of the work and delivered the right material for the target audience, the quality and thoroughness of your work, and presentation including the structure of your review. Remember to spell and grammar check your review before submitting it.

References

[1] Serge Abiteboul, Peter Buneman, and Dan Suciu, *Data on the Web: From Relations to Semistructured Data and XML*, Morgan Kaufmann, 2000.

[2] Michael Barg and Raymond K. Wong, Cooperative Query Answering for Semistructured Data, *Proceedings of the Fourteenth Australasian Database Conference (ADC2003)*.

[3] Ronald Bourret, <http://www.rpbouret.com/xml/>

[4] Ramez Elmasri and Shamkant B. Navathe, *Fundamentals of Database Systems*, Addison-Wesley, 4ed, 2003.

[5] Raghu Ramakrishnan and Johannes Gehrke, *Database Management Systems*, 3ed, McGraw-Hill, 2002.