12.6.3 “Native” Storage

Using “original” concepts of the database for storing XML (internal XML or object model) instead of mapping it or “simply” representing it as ASCII.

- often based on existing object-oriented DB-systems with application of concepts from hierarchical and network-DBs
- no document transformation to another data model
- data model/classes based on the notions of “tree”, “element”, “attribute”, “document order”
- navigation
- XPath/XQuery/XSQL APIs

“Native” Storage: Systems and Products

Many early implementations came from the object-oriented area:

- XML-QL, based on the Strudel system
- LoPiX, based on F-Logic
- Lorel-XML, based on Lorel/OEM
- Tamino (Software AG, Darmstadt, founded 1969 (Adabas, hierarchical DB), Tamino 1999, with XQL, first native XML DBMS),
- Excelon (until 1998: ObjectDesign with “ObjectStore”; since 12.2002: acquired by Progress Software Corp.)
- POET (Hamburg, “Persistent Objects and Extended Database Technology”, product 1990, spin-off from BKS 1992, OQL interfaces, SGML Document Repository 1997, Content Management Suite since 1998, merger with Versant 2004)
- Infonyte (GMD IPSI XQL 1998, based on a “Persistent DOM”, 12.2000: spinoff TU Darmstadt/Fraunhofer-IPSI)

GENERIC DATABASE BEHAVIOR FOR XML DATABASES

Everything that has been developed and dicussed for relational databases is also relevant for XML:

- physical storage + storage management
- optimization, evaluation algorithms
- multiuser operation, transactions (ACID), safety, access control
- ECA-rules, triggers

The algorithms and theoretical foundations are very similar.

Often, relational (or hierarchical) DB technology is actually used inside.

COEXISTENCE OF XML AND RELATIONAL DATA

- generating XML (views, data exchange packets, ...) from stored relational data
- relational (and object-relational) techniques used for efficiently storing data-centric XML
- storing text-oriented data in RDB with specialized “native” datatypes
- XPath is also accepted by SQL/XML
- additional XML processing functionality by packages and object types
- XQuery is still not the “winner” for data-oriented applications!
- is it the winner for document-oriented applications?
  `http://www.w3.org/TR/xquery-full-text/`

Chapter 13
Miscellaneous

Shortcomings in this lecture

- XML: namespaces (used e.g. for XSL, XLink – but in this lecture really interesting only in combination with RDF)
- XML: processing instructions
- theory: XML Data Model, XPath/XQuery Formal Semantics
  (currently in an alignment process with XML Schema)

Application areas:

- Web Services
- Multimedia applications with XML
- Semantic Web

XML Applications (= Languages)

- XHTML
- MathML; see e.g. `http://www.w3.org/Math/XSL/`
- SMIL (Synchronized Multimedia Integration Language): description of multimedia presentations
- MPEG-7: meta-metalevel description of audiovisual contents (i.e., used to describe description languages)
- Web Services (kind of XML-style CORBA successor): WSDL (WS Description Language), WSFL (Web Services Flow Language; Workflows), SOAP (Simple Object Access Protocol – messaging format, not only for Web Services) UDDI (Universal Description, Discovery and Integration); ebXML, MS-BizTalk
- Health Care: HL-7; clinical data exchange
- BIOML (polymer structures), GML (geographic ML), CML (chemical ML)

EPILOGUE

What should have been taught?

- knowledge for practical use of XML
- XML is more than only angle brackets
- the XML world provides examples for many basic concepts of computer science and their combination,
- illustrating how concepts in computer science evolve, and
- ... an idea of developments in the near future:
  the DBIS group is part of it with
  http://dbis.informatik.uni-goettingen.de/rewerse/

13.1 Overview of some Books ...

- There is not a single book that gives a good introduction the everything about XML
- several books on specific, advanced topics
- german books: comments in german.
- recommended sections are marked with $\rightarrow$.
- very recommended sections are marked with $\Rightarrow$.

Note: the selection of books is a bit randomly. There are also other good ones.

“ XML Family of Specifications: A Practical Guide”; Kenneth Sall; Addison-Wesley, May 2002

- historical overview: “professional” focus, no mention of previous research topics
- XML, DTD, SAX; DOM, JDOM; CSS, XSL, XSLT; XLink, XPointer, XPath, XML Schema, RDF, ...
- comprehensive, but often superficial
- 2002 - a bit outdated.

“ XPath, XLink, XPointer and XML”; Erik Wilde, David Lowe; Addison-Wesley, July 2002

- $\Rightarrow$ I recommend the book for its excellent overview of concepts and ideas around XML and the Web (Sections 1-5)
- $\rightarrow$ Sect.6,7: XLink, XPointer
- Sect.8: Usage, Sect.9: Future

“XSLT Programmers Reference”, 2nd Edition; Michael Kay, Wrox Press, June 2003

- $\Rightarrow$ *The* Book on XSLT.

“XQuery”, Wolfgang Lehner and Harald Schöning, dpunkt, 2004

- $\Rightarrow$ *The (german)* Book on XQuery.

“XML & Datenbanken – Konzepte Sprachen und Systeme”; Meike Klettke Holger Meyer; dpunkt-Verlag, 2003

- Kap.3,4 (XML), Kap.10 (Anfragesprachen)
- Kap 2,5,6,7: Allgemeines zu XML und DB
- $\rightarrow$ Kap.8: XML-Datenbanken, Speicherungstechniken
- $\rightarrow$ Kap.9: Indexstrukturen
- $\rightarrow$ Kap.11: XML-Datenbanken: Systeme

“XML und Datenbanken”; Harald Schöning, Hanser-Verlag, 2004

- Kap.1: XML, DTD, DOM, SAX, XSL: oberflächlich
- Kap.2,4: XML Schema, Entwurf
- Kap.3: Allgemeines zum Einsatz von XML
- $\rightarrow$ Kap.5: XML in Datenbanksystemen: allgemeine Betrachtungen
- Kap.6: XPath/XQuery: oberflächlich
- $\Rightarrow$ Kap.7,8: XML und Datenbanksysteme, allgemein sowie Produkte (Oracle, IBM, MS SQL, Tamino)

“XML und Datenmodellierung”; Rainer Eckstein, Silke Eckstein; dpunkt-Verlag, 2004

- grobe Einführung XML, DTD, XPath; (kein XQuery)
- $\Rightarrow$ Kap.4: XML Schema
- (Kap 5: DTD/XML Schema und UML)
- $\Rightarrow$ Kap.6,7: RDF, RFDS, (OWL)