XML as an Interchange Format

Petri Nets 2000

June 27, 2000

Anders Møller

amoeller@brics.dk

BRICS, Department of Computer Science

Aarhus University

Overview

- **What** is XML?
- **Why** use XML?
- **How** to use XML?

What? (1)

Answer #1:

"*XML is a generalization of HTML*"

- **HTML** has a fixed set of markup tags with a fixed semantics
- **XML** has no predefined markup tags – it provides notation for your own markup languages

What? (2)

Answer #2:

"*XML is a replacement for ASCII*"

XML provides:

- Internationalization using Unicode
- Platform independence
- Markup for structuring

What? (3)

Answer #3:

"*XML is simplified SGML*"

SGML: ISO standard from 1985

XML: W3C recommendation from 1998

Canonical XML / **SML**: under development

What? (4)

Answer #4:

"*XML is the future of relational databases*"

XML generalizes the relational data model from **flat tables** to **hierarchical trees**.

SQL is being generalized accordingly.

Why?

You get:

- **Tailor-made markup** for structuring information
- **Internationalization** and **platform independence** for free
- Support by **generic languages and tools**:
  - parsing (DOM, SAX, Expat, XML4J, ...)
  - schema validation (DTD, XML Schema, DSD, ...)
  - transformation (XSLT, ...)
  - querying (XML-QL, ...)
  - linking (XLink, XPointer, XPath)
  - ...

How?

Making your own XML-based Markup Language:

- Define the language syntax
  - use a *schema language* and *validator tools*
  - common approach: *top-down* design
  - consider *modularization* and *reuse*
  - get inspiration from existing XML languages
- Build language-specific supporting tools
  - high-level parsers (on top of XML parsers)
  - transformation stylesheets (for Web publishing)
  - ...
- Agree on using the language
  - support from user organization

How? – Defining the syntax

Using a schema language:

DTD: inherent to XML 1.0, everybody agree that

- it's too complex: horrible syntax and many confusing details
- it's too simple: not enough expressive power

XML Schema: currently a W3C Working Draft

- much larger and more expressive than DTD
- being developed by working group of 52 people
- strong support from industry

DSD: from AT&T and BRICS

- simpler but in many ways more powerful than XML Schema (conditional constraints, context sensitivity, modularization, semi-structured data, linear-time implementation, ...)

How? – Existing languages

- **XHTML** (**www.w3.org/TR/xhtml1**): an XMLization of HTML 4.01
- **CML** (**www.xml-cml.org**): Chemical Markup Language
- **WML** (**www.wapforum.org**): Wireless Markup Language
- **UXF** (**www.yy.cs.keio.ac.jp/~suzuki/project/uxf**):
  UML eXchange Format
- **OMF** (**zowie.metnet.navy.mil/~spawer/JMV-TNG/XML/OMG.html**):
  Weather Observation Markup Format
- **VoiceXML** (**www.voicexml.org**): Voice eXtensible Markup Language
- **ThML** (**www.ccel.org/ThML**): Theological Markup Language

www.oasis-open.org/cover/xml.html#applications
contains a huge list of other serious XML applications

How? – XML for the Web

XLink, **XPointer**, and **XPath** generalize HTML hyperlinks:

- Linking for **general XML**, not just hypertext
- **Multiple sources and destinations** in a single link
- **Relative** addressing
- Addressing of ***points*** and ***ranges***

How? – Using XSLT

XSLT: Extensible Stylesheet Language, Transformations part

- A declarative language based on ***pattern matching*** and ***template instantiation***
- Designed to provide **browser rendering semantics** to XML languages
- Also useful for **general transformation** of XML documents

How? – DOM and SAX

DOM and **SAX**: XML Application Programming Interfaces

- Provide **platform- and language-neutral** standard API for manipulating XML document trees
- Allow easier processing with **general-purpose programming languages**, e.g. Java, C++, JavaScript
- Always an alternative to using domain-specific languages, such as XSLT

More information

- **`www.w3.org/XML`**
  - W3C's XML home page
- **`www.xml.com`**
  - a news site dedicated to XML
- **`www.xmlhack.com`**
  - brief and concise XML news
- **`www.oasis-open.org/cover`**
  - comprehensive XML information
- **`www.brics.dk/~amoeller/XML`**
  - online tutorial on XML and related technologies