SGML is a level of complexity goes in between HTML and future that will transport a lot of contents over the Internet in the future. It is a W3C recommendation. It is a new (1998) markup language.

XML


Reading

2000-02-06

Thomas Krichel

Introducing XML

ICT 80 Lecture 5
XML is more flexible than HTML. It allows to define tags.

XML is less is more flexible than SGML. It does not allow to implement seldom used features of SGML.

XML has not yet replaced HTML, but it may do so one day.

**Design Goals**

1. XML shall be straightforwardly usable over the Internet.
2. XML shall support a wide variety of applications.
3. XML shall be compatible with SGML.
4. It shall be easy to write programs which process XML documents.
5. The number of optional features in XML is to be kept to the absolute minimum, ideally zero.

6. XML documents should be human legible and reasonably clear.
7. The design of XML should be formal and concise.
8. XML documents should be easy to create.
9. Terseness in XML markup is of minimal importance.
element names are case sensitive

each element must have a start and an end tag

elements must be nested

unleft element

the document must have exactly one top level (root) element

well-formed XML documents

```xml
<root

other stuff optional

root element

prolog

XML document structure
```
Comments

Processing instructions

CDATA sections

Character references (discussed in another lecture)

Entity references (discussed in another lecture)

An element may also contain

- #text (use "#text" and \%text to convert text to entity data)
- Character data in elements can contain any data except & (use character entity %text)

Element contents

Do not use the colon in element names.

- The use of the name string "xml" at the beginning of an element name in any capitalization is reserved for future standards.
- Element names must begin with a letter or underscore. They may be followed by zero or more letters, digits, periods, hyphens, and underscores.

Another XML document example
CDATA sections do not nest.

Anywhere where comments can be placed, for example, you can include a HTML page. Can be placed in the CDATA section. You can have any type of character data.

]]> closes a CDATA section.

]]> opens a CDATA section.

CDATA section

Can be placed anywhere where comments can be placed.

Example

<?xml-stylesheet type="text/css" href="styles.css"?>

orge instructions to XML to parse. General form

Processing instructions

Stuff that is between & and

Comments can be placed anywhere except in the markup, like, the

-- I can put & in a comment --

Example:

Start with -- and end with --

Comments
write a one page description of the following XML application

do for next week

In a document Type Declaration or in an XML schema, well-formed, all restrictions that we have discussed apply to every XML document.

Well-formed and valid documents.