

Introduction to Markup Language

Unit 1

What is a markup language?

.A **markup language** is a modern system for annotating a document in a way that is syntactically distinguishable from the text.

.The idea and terminology evolved from the "marking up" of manuscripts, i.e., the revision instructions by editors, traditionally written with a blue pencil on authors' manuscripts.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE recipe PUBLIC "-//Happy-Monkey//DTD RecipeBook//EN"
"http://www.happy-monkey.net/recipebook/recipebook.dtd">
```

```
<recipe>
```

```
  <title>Peanut-butter On A Spoon</title>
```

```
  <ingredientlist>
```

```
    <ingredient>Peanut-butter</ingredient>
```

```
  </ingredientlist>
```

```
  <preparation>
```

```
    Stick a spoon in a jar of peanut-butter,
    scoop and pull out a big glob of peanut-butter.
```

```
  </preparation>
```

```
</recipe>
```

Features

.A common feature of many markup languages is that they intermix the text of a document with markup instructions in the same data stream or file.

```
<html>
<!-- Our first Web page -->

<head>
<title> Markup Languages – Welcome </title>
</head>

<body>

<p> Welcome to Our Web Site! </p>

</body>
</html>
```

Types

.Presentational markup

- The kind of markup used by traditional word-processing systems: binary codes embedded in document text that produce the WYSIWYG effect.

.Procedural markup

- Markup is embedded in text and provides instructions for programs that are to process the text. Well-known examples include troff, LaTeX, and PostScript.

.Descriptive markup

- Markup is used to label parts of the document rather than to provide specific instructions as to how they should be processed.

```
\documentclass[12pt]{article}
\usepackage{amsmath}
\title{\LaTeX}
\date{}
\begin{document}
  \maketitle
  \LaTeX{} is a document preparation system for the \TeX{}
  typesetting program. It offers programmable desktop publishing
  features and extensive facilities for automating most aspects of
  typesetting and desktop publishing, including numbering and
  cross-referencing, tables and figures, page layout, bibliographies,
  and much more. \LaTeX{} was originally written in 1984 by Leslie
  Lamport and has become the dominant method for using \TeX; few
  people write in plain \TeX{} anymore. The current version is
  \LaTeXe.

  % This is a comment; it will not be shown in the final output.
  % The following shows a little of the typesetting power of LaTeX:
  \begin{align}
    E &= mc^2 && \\\
    m &= \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}
  \end{align}
\end{document}
```

LaTeX

LaTeX is a document preparation system for the TeX typesetting program. It offers programmable desktop publishing features and extensive facilities for automating most aspects of typesetting and desktop publishing, including numbering and cross-referencing, tables and figures, page layout, bibliographies, and much more. LaTeX was originally written in 1984 by Leslie Lamport and has become the dominant method for using TeX; few people write in plain TeX anymore. The current version is LaTeX 2_ε.

$$E = mc^2 \quad (1)$$

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2)$$

Well-know markup languages

- HyperText Markup Language (HTML) – the original markup language that was defined as a part of implementing World Wide Web.
- Extensible HyperText Markup Language (XHTML): HTML reformulated in XML syntax.
 - XHTML Basic – a subset of XHTML for simple (typically mobile, handheld) devices. It is meant to replace WML, and C-HTML.
 - XHTML Mobile Profile (XHTML MP) – a standard designed for mobile phones and other resource-constrained devices.
- Mathematical Markup Language (MathML)
- Scalable Vector Graphics (SVG)
- TeX, LaTeX – a format for describing complex type and page layout often used for mathematics, technical, and academic publications.
- Extensible 3D (X3D)

HTML, XHTML and CSS

.Cascading Style Sheets

(CSS) is a style sheet language used for describing the presentation semantics (the look and formatting) of a document written in a markup language. Its most common application is to style web pages written in HTML and XHTML, but the language can also be applied to any kind of XML document.

```
body
{
background-color: #d0e4fe;
}
h1
{
color:orange;
text-align:center;
}
p
{
font-family:"Times New
Roman";
font-size:20px;
}
```

What is XML?

- XML stands for EXtensible Markup Language.
- XML is a markup language much like HTML
- XML was designed to carry data, not to display data.
- XML tags are not predefined. You must define your own tags.
- XML is designed to be self-descriptive
- XML is a W3C Recommendation.

XML DTD

- The purpose of a DTD (Document Type Definition) is to define the legal building blocks of an XML document.
- A DTD defines the document structure with a list of legal elements and attributes.

```
<!DOCTYPE NEWSPAPER [  
    <!ELEMENT NEWSPAPER (ARTICLE+)>  
    <!ELEMENT ARTICLE  
    (HEADLINE,BYLINE,LEAD,BODY,NOTES)>  
    <!ELEMENT HEADLINE (#PCDATA)>  
    <!ELEMENT BYLINE (#PCDATA)>  
    <!ELEMENT LEAD (#PCDATA)>  
    <!ELEMENT BODY (#PCDATA)>  
    <!ELEMENT NOTES (#PCDATA)>  
  
    <!ATTLIST ARTICLE AUTHOR CDATA #REQUIRED>  
    <!ATTLIST ARTICLE EDITOR CDATA #IMPLIED>  
    <!ATTLIST ARTICLE DATE CDATA #IMPLIED>  
    <!ATTLIST ARTICLE EDITION CDATA #IMPLIED>  
>]
```


XML Markup Languages

[.http://en.wikipedia.org/wiki/List_of_XML_markup_languages](http://en.wikipedia.org/wiki/List_of_XML_markup_languages)

- **XHTML**: a markup language that has the same depth of expression as HTML, but with a syntax conforming to XML.
- **XML Schema**: a description of a type of XML document, typically expressed in terms of constraints on the structure and content of documents of that type.
- **XPath**: an expression language for addressing portions of an XML document

XML Markup Languages

[.http://en.wikipedia.org/wiki/List_of_XML_markup_languages](http://en.wikipedia.org/wiki/List_of_XML_markup_languages)

.XQuery: a query language designed to query collections of XML data (similar to SQL).

.XSL Transformations: a language used for the transformation of XML documents.

W3C

.The World Wide Web Consortium (W3C) is an international community where Member organizations, a full-time staff, and the public work together to develop Web standards. Led by Web inventor Tim Berners-Lee and CEO Jeffrey Jaffe, W3C's mission is to lead the Web to its full potential.



Tools

- XML Copy Editor is free software and released under the GNU General Public License.

- Features

- Validation as you type
 - Tag protection
 - XML Schema/Relax NG/DTD support
 - XSLT, XPath

- NotePad++

- Aptna

