XML

X M L

A Computer language which uses mark up (< , >, /, ?, and Character) and is capable of being extended.

What is mark up ?

1. Literally : Notations or Symbols that are used to correct in making up text and indicate how text should be displayed. (<, >, ^, /, ? etc.).
2. Mark-up refers to the sequence of Characters or other Symbols that can be inserted at certain places in a text file, to indicate how a file should look when it is printed or displayed.

What is Extensible ?

1. Literally : Capable of being Stretched out or extended.
2. In Information Technology, extensible describes a program or programming language that is designed so that users and developers can expand or add to its capabilities.

INTRODUCTION TO XML

- XML = eXtensible Markup Language
- Meant to store, transport and exchange data
- Software-independent format for exchanging data
- HTML stores and displays data; XML stores data in a text format that is both human-readable and software-readable/ hardware-readable

Structure of Elements in XML Documents

< root > can be called Parent Element

< child >

< subchild > .... < / subchild >

< / child >

< / root >

XML Declaration

Declaration (Optional)

< ? xml version = “1.0” ? >

(indicates the number of version of xml in use)

or

< ? xml version = “1.0” encoding = “UTF-8” ? >

(indicates the number of version of xml plus character encoding used in document)

UTF stands for UCS Transformation Format ...

UCS means Universal Character SET.

Declaration (Optional)

XML declaration is optional, but if it appears, it must be at the Top. Not even white space or comment should come before it

All XML parsers are required to support “UTF-8” and “UTF – 16” encodings.

XML declaration is Case Sensitive. It should not begin with <?XML…..

XML Tags & Elements

XML elements- logical units of information in an xml document

XML tags- used to mark the start & end of elements

```
<contact-information>
    <name> Samuel Clinton </name>
    <designation> Designer
        <department> Garments  </department>
    </designation>
    <address> 123 Baker Street </address>
    <phone> (123) 456-7890 </phone>
</contact-information>
```

```
<?xml version="1.0" encoding="UTF-8"?> - 1. Declaration

<contact-information>     2. Root Element (Parent) -- (Start Tag)

    <name> Samuel Clinton </name>  --- 3. Child Element
    <designation> Designer  ---  4. Child Element -- (Start Tag)
        <department> Garments </department> -- 5. Sub-Child Element
    </designation> ----  6. (End Tag) of (Line 4) Child Element
    <address> 123 Baker Street </address>  ---- 7. Child Element
    <phone> (123) 456-7890 </phone>  ----  8. Child Element

</contact-information> 9. (End Tag) of (Line 2) Root Element.
```

XML Example – child of message element

```
<?xml version = "1.0" encoding = "UTF-8"?>
<emails>
  <message>
      <to>Alex@example.com</to>
      <from>Brandy@example.com</from>
      <subject>How are you?</subject>
      <body>Hi, let's catch up sometime.</body>
      <attachment></attachment>
  </message>
</emails>
```

Parent of message

Parent of to

Sibling of from

Sibling of to

XML References

References:

A reference is used to add additional text or markup in XML document, where an error may occur if the character is typed directly.

Example : < inside an element -- error

& in text (Bolton & Company) --- error

salary > 2000 --- error

References start with & (ampersand) and end with ; (Semicolon), such as & < > etc.)

References

References start with & and end with ; (& < etc.)

entity references

Type & *for* &
< *for* <
> *for* >
" *for* ”
' *for* ’

character references

Type a for a
&#65 for A
&#90 for z
&# 50 for 2
&# 57 for 9

example : for (Bolton & Company) write (Bolton & Company) . For (Salary>x+1)write (Salary > x+1) etc. The XML parser will change & or > back to & and > automatically when the document is processed.

RULES FOR XML

- **XML is case sensitive**
- **All start tags must have end tags**
- **Elements must be properly nested**
- **XML declaration is the first statement**
- **Every document must contain a root element**
- **Attribute values must have quotation marks**
- **Certain characters are reserved for parsing**

Bookstore

category	title	lang	author	year	price
cooking	Everyday Italian	en	Giada De Laurentiis	2005	30
children	Harry Potter	en	J K. Rowling	2005	29.99
web	Learning XML	en	Erik T. Ray	2003	39.95