CER4C3 Abstractions and Paradigms of Programming

Bachelor of Engineering

Computer Engineering II nd Year IV th Semester

Grammars in XML

- **What is DTD**
- DTD stands for **Document Type Definition**. It defines the legal building blocks of an XML document. It is used to define document structure with a list of legal elements and attributes.
- **Purpose of DTD**
- Its main purpose is to define the structure of an XML document. It contains a list of legal elements and define the structure with the help of them.

Grammars in XML

- A well-formed and valid XML document is one which have been validated against DTD.
- In the previous example we created an XML file for student information. The student information had two elements namely, the name of the student and the branch.
- Now lets say I want to add a new student and while adding that student I forget to enter the branch of the student. There should be some mechanism that warns me whenever I try to insert incomplete information.

Grammars in XML

For the sake of this purpose we create a .DTD file that defines valid syntax covering all the elements and attributes to be included in the XML file.

<students>

<student>

<name>Mahesh</name><branch>CS</branch>

<name>Aashish</name>

</student>

</students>

Grammars in XML

<!ELEMENT students (student)>
<!ELEMENT student(name,branch)>
<!ELEMENT name(#pcdata)>
<!ELEMENT branch(#pcdata)>
<!ATTLIST student sid ID #Required>

Including dtd in XML file
<!DOCTYPE students SYSTEM “StudentType.dtd”>

Including this dtd with XML will validate the data for the presence of all the elements and attributes but will be able to validate only a single element.

Grammars in XML

To have multiple elements inside a single parent tag –

<!ELEMENT students (student+)>

<!ELEMENT student(name,branch)>

<!ELEMENT name(#pcdata)>

<!ELEMENT branch(#pcdata)>

<!ATTLIST student sid ID #Required>

XML Vs HTML

No.	HTML	XML
1)	HTML is used **to display data** and focuses on how data looks.	XML is a software and hardware independent tool used **to transport and store data**. It focuses on what data is.
2)	HTML is a **markup language** itself.	XML provides a **framework to define markup languages**.
3)	HTML is **not case sensitive**.	XML is **case sensitive**.
4)	HTML is a presentation language.	XML is neither a presentation language nor a programming language.
5)	HTML **has its own predefined tags**.	You **can define tags according to your need**.
6)	In HTML, it is **not necessary to use a closing tag**.	XML **makes it mandatory to use a closing tag**.
7)	HTML is **static** because it is used to display data.	XML is **dynamic** because it is used to transport data.
8)	HTML **does not preserve whitespaces**.	XML **preserve whitespaces**.

XML Namespaces

XML namespaces are used to avoid element name conflict in XML document.

- It is a set of unique names.
- Identified by URI (Uniform Resource Identifier)
- Attribute name must start with “xmlns”.

Syntax : - <element xmlns : name = “URI”>

Conflict – Generally conflict occurs when we try to mix XML files from different applications.

Provide Example

XML Namespaces

<students>
<student>
<name>Mahesh</name>
</student>
</students>

<students>
<defaulters>
<name>Aashish</name>
</defaulters>
</students>

Resolving the confusion using namespace

<?xml version=‘1.0’ encoding=‘UTF-8’?>
<info: students xmlns : info=“studentsinformation”>
<info:student>
<info:name>Mahesh</info:name>
</info:student>
</info:students>

<defaulters:students xmlns:defsulters=“studentdefaulters”>
<defaulter:student>
<defaulter:name>Aashish</defaulter:name>
</defaulter:student>
</defaulter:students>

XML Namespaces

- The namespace can be defined by an xmlns attribute in the start tag of an element.
- All child name with the same prefix are associated with the same namespace.