XML PROGRAMMING: SUB CODE- 24662

QPCODE: -780

PART-A

(Each question carries 1 marks, Answer any FIFTEEN(15) questions)

1.Mention any two part of XML tree structure.

- Root node
- Leaf node

2. Write any two use of XPath.
  (Any Two)
- **XPath** is **a** syntax for defining parts of an **XML** document.
- **XPath uses** path expressions to navigate in **XML** documents.
- **XPath** contains **a** library of standard functions.
- **XPath** is **a** major element in **XSLT** and in **XQuery**.
- **XPath** is **a** W3C recommendation.

3. Define WML.

Wireless markup language (WML) is a markup language for wireless devices that adhere to Wireless Application Protocol (WAP) and have limited processing capability.

4.What is an absolute location path?

Location path specifies the location of node in XML document. This path can be absolute or relative. If location path starts with root node or with '/' then it is an absolute path. Following are few of the example locating the elements using absolute path.

5.Write the limitation of schema language.

- Poor support for xml namespace
- Poor data typing
- Limited content model description
- It supports only the text string data type.
- Limited possibilities to express the cardinality for elements.

6.Mention any two declaration that can be used in DTDs.
  (Any Two)

- <!ELEMENT letter (date, address, salutation, body, closing, signature)>
        Element – letter
        Child Elements - date, address, salutation, body,closing,signature

- <!ELEMENT Name (#PCDATA)>
        #PCDATA – parsed character data
            - Data contains only text

- <!ELEMENT Street (#CDATA)>
    #CDATA – Character Data
      - Data may contain text, numbers and other character.
- <!ELEMENT br EMPTY>
    EMPTY – It has no content.
- <!ELEMENT body(p+)>
    p+ - < body> element may contain more than one <p> elements.

7.What is a namespace?

Namespaces provide a method to avoid element name conflicts.

8.Define validation.

Validation is a process by which an XML document is validated. An XML document is said to be valid if its contents match with the elements, attributes and associated document type declaration (DTD), and if the document complies with the constraints expressed in it.

Validation is done on

- DTD validation
- Schema validation

9.What is XSLT?

XSLT (eXtensible Stylesheet Language Transformations

XSLT is far more sophisticated than CSS. With XSLT you can add/remove elements and attributes to or from the output file.

10.What is the statement used for sorting?

The statement used for sorting is
    <XSL: sort select =" expression">
    Order=" ascending/descending"

11.Which instruction cause the text of the title to be processed?

<t xml: lang="en">

This is some text contained by the 't' element. The use of the xml:lang attribute indicates the language so that, for example, the correct font could be applied when rendered or the correct spell-checker could be used when proofing the document.

12.Define multiple documents.

The multiple documents to the database collection by using a single command. To do so, you use add_multiple_documents (alias add multiple) on the xindice utility.

A multiple document in this manner, xindice creates a unique key for each added document. This key has a value that is the same as its filename.

13.What is an arbitrary tree.

Arbitrary tree is XML documents form a tree structure that starts at "the root" and branches to the "leaves".

14.Expand: AJAX.

AJAX = **A**synchronous **J**avaScript **a**nd **X**ML.

15.What is the data type expression?

The Expression data type allows you to express an arbitrary formula as a string, and instantly evaluate it.

16.Define prolog.

Prolog is the component added in the beginning of an XML document.

Prolog includes XML declaration, DOCTYPE and comments, processing instructions too.

17.What is DOM?

The Document Object Model is a cross-platform and language-independent interface that treats an XML or HTML document as a tree structure wherein each node is an object representing a part of the document. The DOM represents a document with a logical tree.

18.Which organization does standardize DOM?

- World wide web consortium
- WHATWG (web hypertext application technology working group)

19.What is the STXPath?

STX is a variation of XSLT suitable for streaming
- Some features are not allowed
- But every stx application can be streamed.
- The differences reflect necessary limitations in the control flow.

20. Expand XACT.

- ❖ XACT stands for – X Audio Compression Toolkit

PART-B

21. a). (i) Describe the general form of location steps and explain the parts of it with an example. [6]

A Location step consists of three components –
1. An axis
2. A node test and
3. A predicate

General form: axis: node test[predicate]

If axis is present: is required.

If a predicate is present, [] is required

The purpose of these three components is used to select a node that bears some relationship to the context node.

The Axis: The axis tells the processor which direction to look from the context node. The difference axes supported by XPath processors are
- self- locates the context node itself
- attribute – locates an attribute of the context node.
- child – locates a node immediately descended from the context node
- descendant – locates all nodes descendant from context node etc

Node test: The node test portion of location steps narrows the view of all the nodes available along the given axis to a subset of them. It can take any three possible form :
- node type – to search along the specified access for nodes of a particular type
- name – name of a node to be tested
- specified PI –to locate the specific processing instruction using the following form
      processing-instruction (target)

Predicate: The predicate portion of the XPath expression refined the set of nodes established by the axis and node test.

Example: - name (/node () [2])

(ii) Write any six XPath axis names and explain. - [6]

(Any six)
- self- locates the context node itself
- attribute – locates an attribute of the context node.
- child – locates a node immediately descended from the context node
- descendant – locates all nodes descendant from context node etc
- descendant-or-self – locates all descendants of the context node as well as context node itself
- ancestor – locate all elements in a direct line above the context node
- ancestor-or-self-locate all ancestors of the context node as well as context node itself
- preceding - locate all node before the context node
- preceding-sibling - locate all preceding nodes that share the same parent as the context node.
- following- locate all node after the context node
- following-sibling- locate all following node that share the same parent as the context node.
- namespace- locate all the namespace nodes in effect for the contest node

(OR)

B) List and explain the different forms of expressions. [12]

Every expression evaluates to a sequence of atomic values or nodes. The different forms of expression are.

(i) Literal expression
(ii) comment expression
(iii) Arithmetic expression
(iv) sequence expression
(v) Path expression
(vi) Filter expression
(vii) Comparison expression
(viii) Boolean expression.
(ix) conditional expression
(x) Quantified expression.

i) Literal expression :

Literal expression denotes a constant atomic value. for example,

→ string literal are written either single or double quotation marks. Ex: "Paris" (or) 'Paris'.

→ Integer literal are written as 42.

→ Decimal literal are written as 3.1412.

→ Negative number can be written as -93.7 and minus sign is not part of literal.

(ii) Comment expression :

XPath 2.0 expression may be commented using this syntax:

(: comment :)

Ex:

(: multiply each day by two :) /palanets/ planets / day / (for $i in. return $ i*2).

(iii) Arithmetic expression :

→ The arithmetic operator are :

\+ , − , * , div , mod, idiv.

→ The idiv operates only on integers and returns an integer rounded toward zero. The other operator operates all four kind of numbers.

Ex: 3 idiv 2   returns 1.

−5 idiv 2   returns −2.

(iv) Path expression :

→ XPath uses path expressions of select nodes or node sets in an XML document.

nodename	description.
/	Selects all nodes with the name.
.	selects current node
@	select attributes.

Ex:

(v) Boolean Expression:

→ is Boolean can only have two values : true or false.

→ The boolean operators are : and , or

Ex: count =0   or   count =1

22) a) (i) Describe about modularization. - [6]

      Modularization is a framework for defining markup languages.

      These schemas define the structure of content developed using the languages in addition to defining the elements and attributes that make up the vocabulary of the languages. Such markup languages are orthogonal to the concept of "XML Namespaces".

      The first and the most important document in this family is "Modularization of XHTML". This document defines the way in which XHTML-family modules are defined at the abstract level, and decomposes "XHTML 1.0" into sets of related elements and attributes.

      These sets are called modules. They are designed such that there are a few core modules. These must be present in all XHTML-family document types. These core modules provide structure for the document type as well as simple text markup.

      These minimal requirements for XHTML-family document types are designed so to ensure that XHTML conforming documents will render on all XHTML-conforming systems, regardless of the XHTML document type to which they are written.

(ii) Explain briefly about DTD with example. – [6]

      DTD Stands for "Document Type Definition."
      A DTD defines the structure and the legal elements and attributes of an XML document. With a DTD, independent groups of people can agree on a standard DTD for interchanging data. An application can use a DTD to verify that XML data is valid.

Seen from a DTD point of view, all XML documents are made up by the following building blocks:

- Elements
- Attributes
- Entities
- PCDATA
- CDATA

Example:

Newspaper Article DTD

<!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>

```
<!ENTITY NEWSPAPER "Vervet Logic Times">
<!ENTITY PUBLISHER "Vervet Logic Press">
<!ENTITY COPYRIGHT "Copyright 1998 Vervet Logic Press">
>
```

(OR)

b. Tabulate the XML schema data types and describe them with examples. - [12]

Name	Description
date -Ex: <date>2019-10-15</date>	Defines a date value
dateTime Ex: <date><time>2019-10-15:10:25:</date></time>	Defines a date and time value
Duration Ex: <starttime><startdate>15:10:25:…2012:09:26</starttime></startdate>	Defines a time interval
gDay Eg: <gday>…30</gday>	Defines a part of a date - the day (DD)
gMonth<gmonth>..12</gmonth>	Defines a part of a date - the month (MM)
gYear Ex:<gyear>…2000</gyear>	Defines a part of a date - the year (YYYY)
gMonthDay Ex:<gmonthday>..00-12</gmonthday>	Defines a part of a date - the month and day (MM-DD)
gYearMonth-Ex:<gyearmonth>1998-02</gyearmonth>	Defines a part of a date - the year and month (YYYY-MM)
Time Ex:<starttime>10:20:20</starttime>	Defines a time value

23) a) (i) Draw the processing model of XSLT processing and explain. - [6]

Processing model of XSLT.

- When an XSLT processor applies a stylesheet to a target document.
- It finds a template rule that is appropriate for the document root and execute the corresponding template body.

XML Input   XSLT Code

XSLT Processor

Result Document
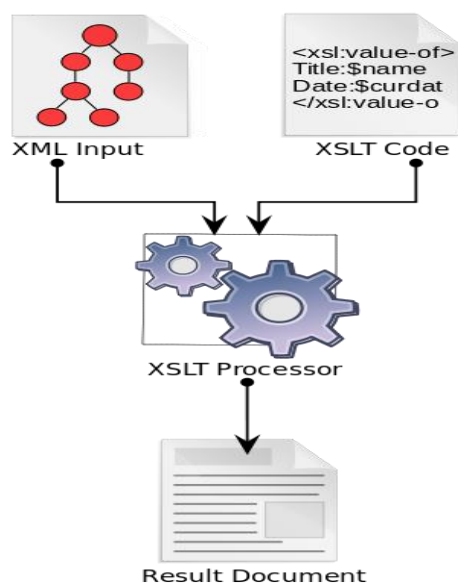
      The XSLT processor takes one or more XML source documents, plus one or more XSLT stylesheets, and processes them to produce an output document. In contrast to widely implemented imperative programming languages like C, XSLT is declarative. The basic processing paradigm is pattern matching. Rather than listing an imperative sequence of actions to perform in a stateful environment, template rules only define how to handle a node matching a particular XPath-like pattern, if the processor should happen to encounter one, and the contents of the templates effectively comprise functional expressions that directly represent their evaluated form: the result tree, which is the basis of the processor's output.

(ii) Write the format of XSLT style sheet and explain. - [6]

      <?XML version="1.0" encoding =" UTF-8"?>

      <XSL: stylesheet version="1.0"

      Xmlns:xsl=http://www.w3.org/1999/xsl/Transform>

      <xsl: template match=" ….">

      ……….

      </XSL: stylesheet>

      <XSL: stylesheet>Element:

A stylesheet is a collection of rules.

The <XSL: stylesheet> and <XSL: transform>

Elements are completely identical elements.

Both are used to define the root element of the style sheet.

Syntax:

<XSL: stylesheet version="1.0"

Xmlns:xsl=http://www.w3.org/1999/xsl/Transform>

(OR)

b. What are the two major kinds of template rules? Explain the built-in template rules of the seven types of nodes. - [12]

A template specifies the instructions for the transformation
There are two components:

- Identifying the nodes in the XML document that are used in the transformation
- Details about the actual transformation

There are **seven kinds of nodes** in an XML document:

- the root node, - the **root node** is the very first or parent **node**
- element nodes, - element refers to an element or a node of type element returned by a member of the DOM API.
- attribute nodes, - When an attribute is returned by a member, it is an object reference that exposes a special (albeit small) interface for attributes.

- text nodes, - Represents textual content in an element or attribute
- comment nodes, - Represents a comment
- processing instruction nodes, - Represents a processing instruction
- namespace nodes. - A namedNodeMap is like an array, but the items are accessed by name or index

24) a) Explain in detail about XML data bases. – [12]

XML Database is used to store the huge amount of information in the XML format. As the use of XML is increasing in every field, it is required to have the secured place to store the XML documents. The data stored in the database can be queried using XQuery, serialized, and exported into desired format.

There are two major types of XML databases:
- XML- enabled
- Native XML NXD

XML- Enabled Database XML enabled database is nothing but the extension provided for the conversion of XML document. This is relational database, where data are stored in tables consisting of rows and columns. The tables contain set of records, which in turn consist of fields.

Native XML database is based on the container rather than table format. It can store large amount of XML document and data. Native XML database is queried by the XPath-expressions.

Native XML database has advantage over the XML-enabled database. It is highly capable to store, query and maintain the XML document than XML-enabled database.

Example:

```xml
<?xml version="1.0"?>
<contact-info>
    <contact1>
        <name>Tanmay Patil</name>
        <company>TutorialsPoint</company>
        <phone>(011) 123-4567</phone>
    </contact1>
    <contact2>
        <name>Manisha Patil</name>
        <company>TutorialsPoint</company>
        <phone>(011) 789-4567</phone>
    </contact2>
</contact-info>
```

(OR)

b) Discuss in detail about AJAX and its components. – [12]

AJAX stands for **A**synchronous **Ja**vaScript and **X**ML. AJAX is a new technique for creating better, faster, and more interactive web applications with the help of XML, HTML, CSS, and Java Script.

- Ajax uses XHTML for content, CSS for presentation, along with Document Object Model and JavaScript for dynamic content display.

- Conventional web applications transmit information to and from the sever using synchronous requests. It means you fill out a form, hit submit, and get directed to a new page with new information from the server.

- With AJAX, when you hit submit, JavaScript will make a request to the server, interpret the results, and update the current screen. In the purest sense, the user would never know that anything was even transmitted to the server.

- XML is commonly used as the format for receiving server data, although any format, including plain text, can be used.

- AJAX is a web browser technology independent of web server software.

- A user can continue to use the application while the client program requests information from the server in the background.

- Intuitive and natural user interaction. Clicking is not required, mouse movement is a sufficient event trigger.

- Data-driven as opposed to page-driven.

The key technical components of AJAX are:

- XHTML – a stricter, cleaner rendering of HTML into XML.

JavaScript

- Loosely typed scripting language.
- JavaScript function is called when an event occurs in a page.

11

- Glue for the whole AJAX operation.

DOM

- API for accessing and manipulating structured documents.
- Represents the structure of XML and HTML documents.

CSS

- Allows for a clear separation of the presentation style from the content and may be changed programmatically by JavaScript

XMLHttpRequest

- JavaScript object that performs asynchronous interaction with the server.

25) a) Explain about XSLT transformation. - [12]

XSLT is a language for transforming XML documents. the XSLT processor is concerned with three XPath data model trees: the source tree, the stylesheet tree, and the result tree. Fig. shows the relationship between these three. The stylesheet and source trees are fed to the XSLT processor, which then produces the result tree.
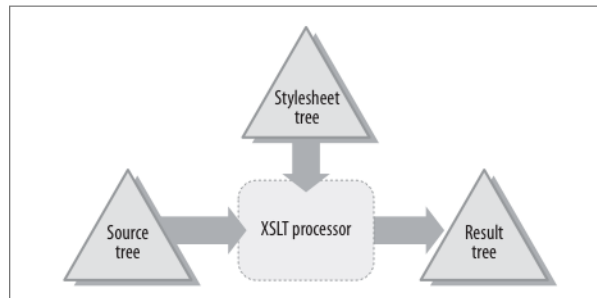
The tree consists of nodes that can be the following:

- Root, which can include the root and all that it contains
  The root node is different from the root element in that it contains the whole document, including all the child elements, comments and processing instructions
- Elements
- Attributes
- Text, the value of the element between the tags
- Namespaces
- Processing instructions
- Comments

XSLT is a *declarative language*, providing a series of templates for defining the transformation
The templates match patterns (usually XML elements) which consist of both XML elements, attributes and other XML nodes.
Templates do not usually contain programming code in the usual sense, such as variables (although they may include variables and IF statements).

XSLT Stylesheet format

The XSLT document is an XML document and must begin with two declarations

<?xml version="1.0"?>
<xsl: stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">

and end with

</xsl: stylesheet>

You can also use the *synonym* for stylesheet, the xsl: transform. This replaces the word stylesheet with the word transform:

<?xml version="1.0"?>
<xsl: transform version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">

and end with

</xsl: transform>

Stylesheet guidelines

- An XSLT stylesheet contains template rules, which specifies the nodes (the elements, etc. to be used), the input and the output
- The XSLT document is an XML document and *must* be well-formed
- If an element of the XSL template is *not* an xsl rule, it is not transformed
- These are literals and may be any text - they are output exactly as they are defined
- If the literals are HTML code, they must conform to XML rules (you must use xhtml)

- This allows you to include HTML elements, etc. in the document

Example (creating a delimited file - suitable for importing data into a database)
Stylesheet

(OR)

b) (i) Explain about SAX filters. - [6]

A SAX filter is simply a class that is passed as the event handler to another class that generates SAX events, then forwards all or some of those events on the next handler (or filter) in the processing chain. A filter may prune the document tree by not forwarding events for elements with a given name (or that meet some other condition), while in other cases, a filter might generate its own new events to add parent or child elements to certain elements the existing document stream.

Also, element attributes can be added or removed or the character data altered in some way. Really any class that is able to receive SAX events, then call event methods on another SAX handler in a way that alters the document stream can be seen as a SAX filter.
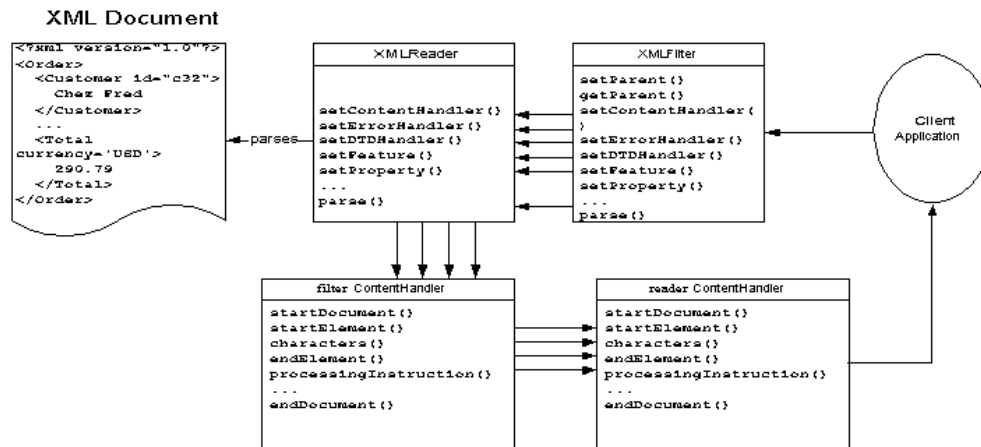
XML Document

This step should be done at the time of using filter

- Create a filter object. Invoking by its own constrictor.
- Create the XMLReader that will pass the document **by** calling XMLReaderFactory.createXMLReader().
- Attached the filter to parser using the filter stepparent () method.
- Install the content Handler in the filter
- parse the document by calling filter's parse () method.

(ii) Write the limitations of streaming. - [6]

- ❖ In theory the expressive power of streaming and the in-memory concept is the same.
- ❖ The programmer is responsible not to blow up the streaming concept.
- ❖ Assumption: memory is limited
    - In STX significant rearrangements are not possible.
    - Sorting is (almost) impossible.
    - For every size of main memory one can construct an XML document, which is not processable in STX.

Date   :   Prepared by
Location   : CHENNAI   **MALARVIZHI K [41205206]**
                                                        Part Time / Guest Lecturer
                                          120,   Government   Polytechnic
College,
                                                   Purasaiwakkam,
                                                   Chennai – 12.