S. Aiwarya
310620104003

16.2.23

SERVICE ORIENTED ARCHITECTURE

ASSIGNMENT - 1

1. **SAX:**

* Simple API for XML is an event-driven online algorithm for lexing and parsing XML documents, with an API developed by the XML-DEV mailing list.

* SAX is a way of reading data from an XML document that is alternative to the Document Object Model's mechanism.

* DOM works on the document as a whole, creating the whole abstract syntax tree of XML document for user's convenience whereas SAX parsers work on each element of XML document sequentially, issuing parsing events.

* Java implementation of SAX is regarded as the de-facto standard.

* Parsers are used to process XML documents where the parser examines the document, check for errors and then validate it against DTD if it is a validating parser.

* Apart from triggering certain events, SAX parser does nothing with the data.

* The SAX events are:

    i) XML Text Nodes
    ii) XML Element Starts and Ends
    iii) XML Processing Instructions
    iv) XML Comments.

* There are no update methods in SAX parser and since the complete file is not kept in memory, it is possible to access items only in a sequential manner.

* SAX parser is used when

    i) XML document can be processed in a linear fashion
    ii) The document is not deeply nested.
    iii) The problem to be solved involves only a part of the XML document.
    iv) SAX works well for an XML document that arrives over a stream.

CONTENT HANDLER INTERFACE:

* This interface specifies callback methods that SAX parser uses to notify an application program of the components.

i) void startDocument ()

ii) void endDocument ()

iii) void startElement (String uri, String localName, String qName, Attributes atts)

iv) void setDocumentLocator (Locator locator)

v) void startPrefixMapping (String prefix, String uri)

vi) void endPrefixMapping (String prefix)

2. XPATH:

* XPath stands for XML Path Language.

* It is a major element in the XSLT standard and W3C recommendation.

* XPath uses path expressions to select nodes or node-sets in an XML document.

* XPath includes functions for string values, numeric values, booleans, date and time comparison, node manipulation, sequence manipulation etc.

* Some of the features of XPath are:

　i) It defines structure

　ii) It provides path expression

　iii) It is a standard function and it is W3C recommendation

* Some of the syntax of XPath expressions are:

  i) nodename - Selects all nodes with name "nodename"

  ii) / - Selects from rootnode

  iii) // - Selects nodes in document from current node

  iv) . - Selects currentnode.

  v) .. - Selects parent of currentnode.

  vi) @ - Selects attributes.

* Ex:
```
<?xml version="1.0" encoding="UTF-8"?>
<bookstore>
    <book>
    <title lang="en"> ABC </title>
    <price> 110 </price>
    </book>
</bookstore>
```

  i) /bookstore - Selects root element bookstore.

  ii) bookstore/book - Selects all book elements.

  iii) //book - Selects all book elements no matter where they are in document.

  iv) //@lang - Selects all attributes that are named lang.

* Wildcards are used to select unknown XML nodes.
* Predicates are used to find specific node or node that contains a specific value.

3. XSL :

* XSL is a family of standards that specifies how to define XML document transformation and presentation.

* It stands for Extensible Stylesheet Language.

* XSL was a single standard that is used to display an XML document of a given type, similar to CSS.

* Initially XSL was divided into 3 parts:

      i) XSLT (XSL Transformations)

      ii) XPath

      iii) XSL-FO (XSL Formatting Objects)

i) XSLT :

* XSLT standard defines a language for transforming

* XML documents into other XML documents or into other formats such as HTML, Extensible HTML, text or JSON (JavaScript Object Notation).

* An XSLT document is essentially a stylesheet written in

well- formed XML to provide specific styling information for one or more XML source documents.

* To use XSLT, XSLT processor is needed that takes XML and XSLT documents as input, parses XML, applies the XSLT to the parsed data and generates an output document in the specified format such as XML, HTML or text.

ii) XPath:

* It defines compact expression languages used by XSLT and other languages to access parts of an XML document.

* Its primary purpose is to address the nodes in XML trees.

iii) XSL-FO:

* It defines a template based vocabulary and expression language for specifying XML formatting semantics.

* It provides a type of markup language.

4. XQUERY:

* XQuery is a functional query language used to retrieve information stored in XML document.

* It is built on XPath expressions.

* A standardized language for combining documents, databases, web pages and almost anything else.

* XQuery is replacing Web Application development languages.

* It is used to:

    i) Extract information to use in a web service.

    ii) Generate summary reports

    iii) Transform XML to HTML.

    iv) Search web documents for relevant information.

* Some characteristics of XQuery are:

    i) Functional language

    ii) XPath based

    iii) Universally accepted

    iv) W3C standard.