XML and DTD

XML

XML stands for extensible Markup Language

1. XML is designed to store and transport data

2. XML is designed to be self-descriptive

3. XML is a markup language like HTML

4. XML Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format.

XML

- Similar to HTML, but without predefined tags to use. Instead, you define your own tags designed specifically for your needs. This is a powerful way to store data in a format that can be stored, searched, and shared.

- Since the fundamental format of XML is standardized, if you share or transmit XML across systems or platforms, either locally or over the internet, the recipient can still parse the data due to the standardized XML syntax.

Design Rules

- For an XML document to be correct, the following conditions must

  be fulfilled:
- Document must be well-formed.
- Document must conform to all XML syntax rules.
- Document must conform to semantic rules, which are usually set in an XML schema.

Sample Code

```
<?xml version="1.0" encoding="UTF-8"?>
<message>
   <warning>
      Hello World
   </warning>
</message>
```

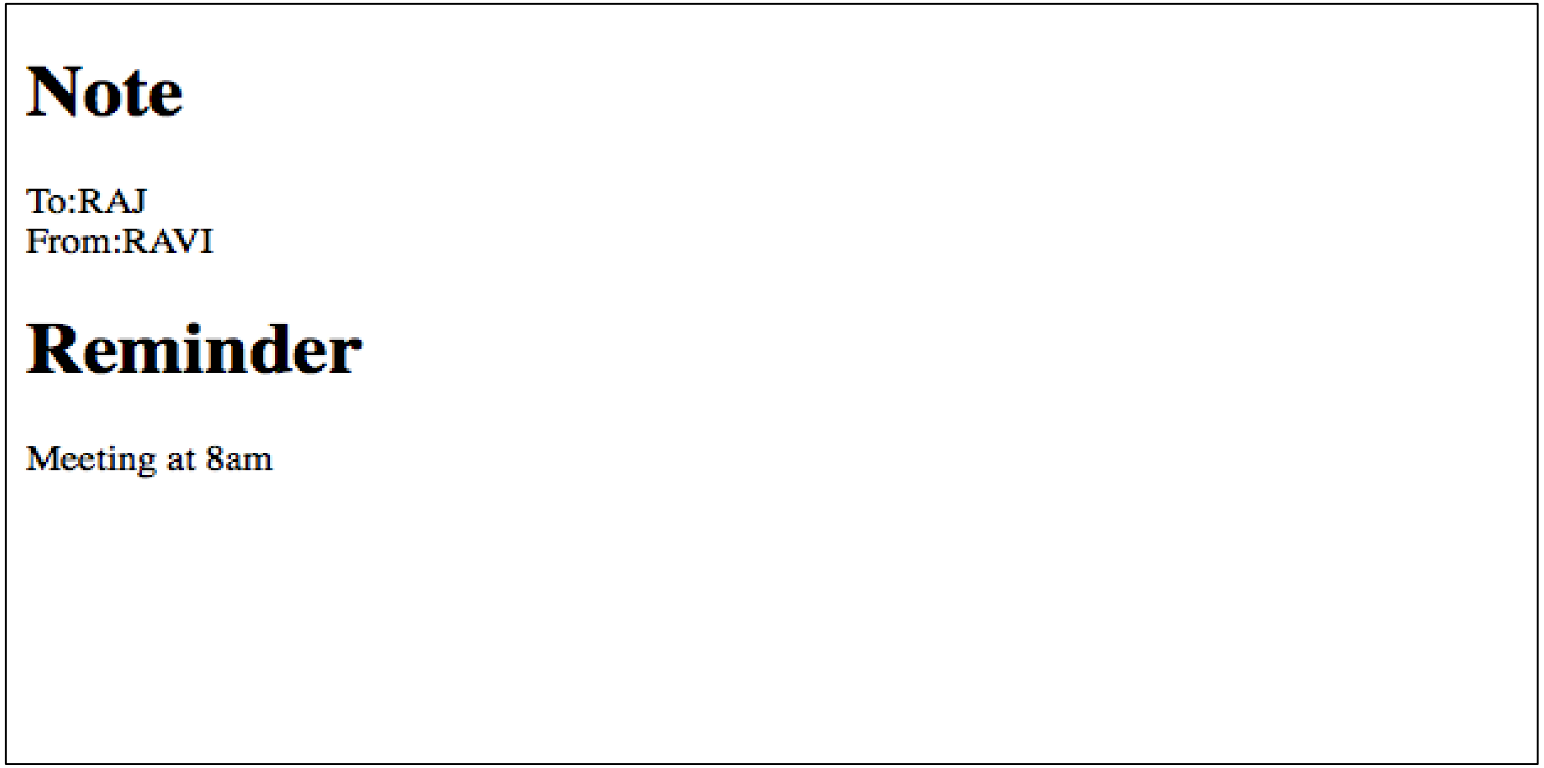
- XML - declaration is not a tag.
- This line is called XML Prolog or XML declaration.
- This line is optional i.e, it can be either used or not in an XML document. However, it should be the very first line if used.

Attributes

version: Used version XML in this document.

encoding: Used encoding in this document.

Output



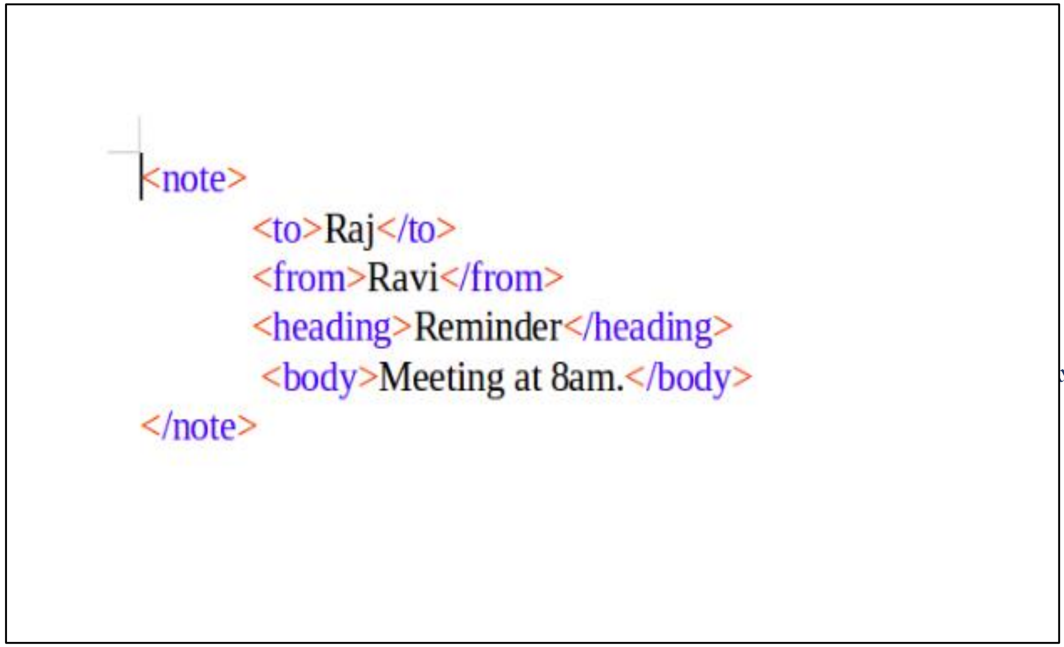
Note

To:RAJ
From:RAVI

Reminder

Meeting at 8am

XML CODE/HTML CODE



```
<note>
          <to>Raj</to>
          <from>Ravi</from>
          <heading>Reminder</heading>
           <body>Meeting at 8am.</body>
</note>
```

```
<html>
<h1>Note</h1>
<body>
<p>To:RAJ
<br>
From:RAVI
</p>
<h1>Reminder</h1>
<p>Meeting at 8am</p>
</body> </html>
```

The output in both the cases is same but while using HTML we have used predefined tags like p tag and h1 tag whereas while using XML we have used self defined tags like “To” tag and “from” tag. The tags in the example are not defined in any XML standard. These tags are “invented” by the author of the XML document

XML –no root element

```
to>Students</to>
<from>Teacher</from>
<subject>Regarding assignment
submission</subject>
<text>All students will have to
submit assignment by
tomorrow.</text>
```

```
<?xml version="1.0"
encoding="UTF-8"?>
<message>
<to>Students</to>
<from>Teacher</from>
<subject>Regarding assignment
submission</subject>
<text>All students will have to
submit assignment by
tomorrow.</text>
</message>
```

Syntax rules for XML:

- In the first example the Root element is <message> and all the remaining elements <to>, <from> etc is the child element and reside within the root element.

- It is case sensitive.

- The XML element should have a closing element for example **<text category = “message”>Hi</text>** is correct but **<text category = “message”>Hi** is not correct because it does not contain the closing element and it will throw an error and vice-versa.

- The elements in XML should be nested properly otherwise it will throw an error. For example **<to><from>Geeks</from></to>** is nested correctly but **<to><from>Geeks</to></from>** is wrong because if <from> is opened inside the <to> element then this should also end inside of the </to> element.

Syntax rules for XML:

- The XML attribute is having two part one is Name and other is its value. It resides inside of the opening of an XML element. For example: <text category = “message”>All students will have to submit the assignment by tomorrow.</text> Here category is the attribute name and message is its value and the attribute value should either be in a single quotation or in double quotation otherwise it will throw an error. The Attribute Name is written without any quotation.

- An XML element can have multiple attributes but can not have the same attribute names in the same element. For example: <text category =”message” purpose = “greet”>GeeksforGeeks</text> Above attributes is correct because of having multiple attributes with the different attribute name. <text category =”message” category = “greet”>GeeksforGeeks</text> Above attribute is wrong because of having the same attribute name in a single element.

Sample Code for Displaying Books.xml

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

<?xml-stylesheet type="text/css" href="Rule.css"?>

<books>

   <heading>Welcome To GeeksforGeeks </heading>

   <book>

      <title>Title -: Web Programming</title>      <author>Author -: Chrisbates</author>      <publisher>Publisher -: Wiley</publisher>

      <edition>Edition -: 3</edition>

      <price> Price -: 300</price>    </book>

   <book>

      <title>Title -: Internet world-wide-web</title>      <author>Author -: Ditel</author>      <publisher>Publisher -: Pearson</publisher>

      <edition>Edition -: 3</edition>

      <price>Price -: 400</price>    </book>

   </books>
```

Rule.css

```
books {
    color: white;
    background-color : gray;
    width: 100%;   }
 heading {
    color: green;
    font-size : 40px;
    background-color : powderblue;   }
 heading, title, author, publisher, edition, price {
    display : block;   }
 title {
    font-size : 25px;
    font-weight : bold;}
```