Module – 1
Introduction to HTML

1.1 A Brief Introduction to the Internet,
1.2 WWW,
1.3 Web Browsers and Web Servers,
1.4 URLs,
1.5 MIME,
1.6 HTTP,
1.7 Security,
1.8 The Web Programmers Toolbox.
1.9 XHTML: Basic syntax,
1.10 Standard structure,
1.11 Basic text markup,
1.12 Images,
1.13 Hypertext Links.

1.1 A BRIEF INTRODUCTION ABOUT THE INTERNET

1.1.1 Origins:

- **1960s**
  - U.S. Department of Defence (DoD) became interested in developing a new large-scale computer network
  - The purposes of this network were communications, program sharing, and remote computer access for researchers working on defence-related contracts.
  - The DoD's Advanced Research Projects Agency (ARPA) funded the construction of the first such network. Hence it was named as ARPAnet.
  - The primary early use of ARPAnet was simple text-based communications through e-mail.
- **late 1970s and early 1980s**
  - BITNET, which is an acronym for ***B**ecause **I**t's **T**ime **NET**work*, began at the City University of New York. It was built initially to provide electronic mail and file transfers.
  - CSNET is an acronym for ***C**omputer **S**cience **NET**work*. Its initial purpose was to provide electronic mail.
- **1990s**
  - NSFnet which was created in 1986 replaced ARPAnet by 1990.
  - It was sponsored by the National Science Foundation (NSF).
  - By 1992 NSFnet, connected more than 1 million computers around the world.

- In 1995, a small part of NSFnet returned to being a research network. The rest became known as the *Internet.*

1.1.2 What the Internet is:

- The Internet is a huge collection of computers connected in a communications network.
- The Transmission Control Protocol/Internet Protocol (TCP/IP) became the standard for computer network connections in 1982.
- Rather than connecting every computer on the Internet directly to every other computer on the Internet, normally the individual computers in an organization are connected to each other in a local network. One node on this local network is physically connected to the Internet.
- So, the Internet is actually a *network of networks*, rather than a network of computers.
- Obviously, all devices connected to the Internet must be uniquely identifiable.

1.1.3 Internet Protocols (IP) Addresses

- The Internet Protocol (IP) address of a machine connected to the Internet is a unique 32-bit number.
- IP addresses usually are written (and thought of) as four 8-bit numbers, separated by periods.
- The four parts are separately used by Internet-routing computers to decide where a message must go next to get to its destination.
- Although people nearly always type domain names into their browsers, the IP works just as well.
- For example, the IP for United Airlines (www.ual.com) is 209.87.113.93. So, if a browser is pointed at http://209.87.113.93, it will be connected to the United Airlines Web site.

1.1.4 Domain names

The IP addresses are numbers. Hence, it would be difficult for the users to remember IP address. To solve this problem, text based names were introduced. These are technically known as ***domain name system (DNS).***

These names begin with the names of the host machine, followed by progressively larger enclosing collection of machines, called ***domains.*** There may be two, three or more domain names. DNS is of the form **hostname.domainName.domainName** . Example:**atme.ac.in** The steps for conversion from DNS to IP:

- The DNS has to be converted to IP address before destination is reached.
- This conversion is needed because computer understands only numbers.
- The conversion is done with the help of *name server*.
- As soon as domain name is provided, it will be sent across the internet to contact name servers.
- This name server is responsible for converting domain name to IP
- If one of the *name servers* is not able to convert DNS to IP, it contacts other name server.
- This process continues until IP address is generated.

- Once the IP address is generated, the host can be accessed.
- The hostname and all domain names form what is known as FULLY QUALIFIED DOMAIN NAME.

This is as shown below:

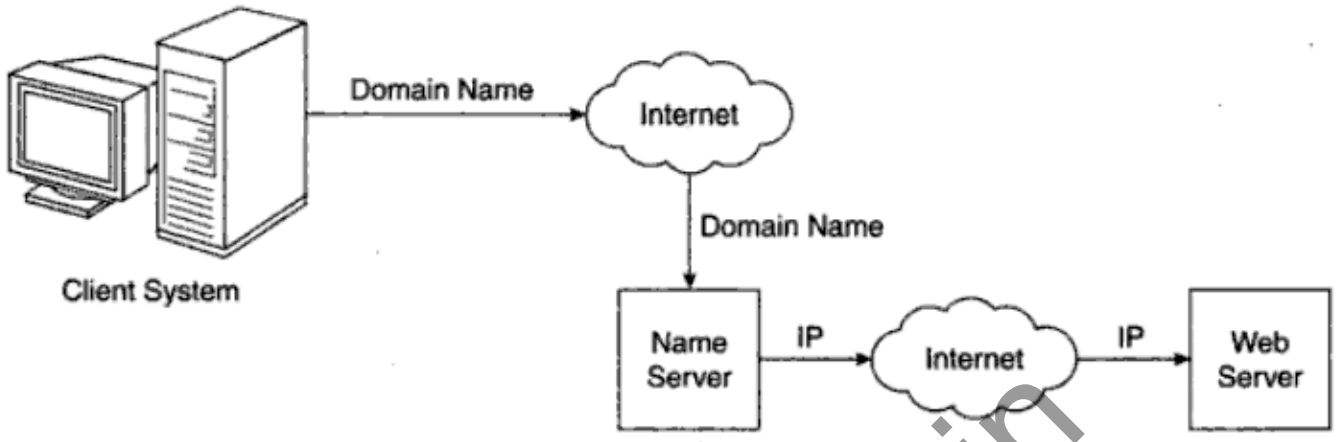


Figure 1.1 Domain name conversion

1.2 The World-Wide Web

1.2.1 Origins

- Tim Berners Lee and his group proposed a new protocol for the Internet whose intention was to allow scientists around the world to use the Internet to exchange documents describing their work.
- The proposed new system was designed to allow a user anywhere on the Internet to search for and retrieve documents from the databases on any number of different document-serving computers.
- The system used *hypertext*, which is text with embedded links to text in other documents to allow non-sequential browsing of textual material.
- The units of web are referred as pages, documents and resources.
- Web is merely a vast collection of documents, some of which are connected by links.
- These documents can be accessed by web browsers and are provided by web servers.

1.2.2 Web or Internet?

It is important to understand that the Internet and the Web is not the same thing.

- The **Internet** is a collection of computers and other devices connected by equipment that allows them to communicate with each other.
- The **Web** is a collection of software and protocols that has been installed on most, if not all, of the computers on the Internet.

1.3 Web Browsers

- Documents provided by servers on the Web are requested by **browsers**, which are programs running on client machines.

- They are called browsers because they allow the user to browse the resources available on servers.
- Mosaic was the first browser with a graphical user interface.
- A browser is a client on the Web because it initiates the communication with a server, which waits for a request from the client before doing anything.
- In the simplest case, a browser requests a static document from a server.
- The server locates the document among its servable documents and sends it to the browser, which displays it for the user.
- Sometimes a browser directly requests the execution of a program stored on the server. The output of the program is then returned to the browser.
- Examples: Internet Explorer, Mozilla Firefox, Netscape Navigator, Google Chrome, Opera etc.,

1.4 Web Servers

Web servers are programs that provide documents to requesting browsers. Example: Apache

1.4.1 Web Server Operation:

- All the communications between a web client and a web server use the HTTP
- When a web server begins execution, it informs the OS under which it is running & it runs as a background process
- A web client or browser, opens a network connection to a web server, sends information requests and possibly data to the server, receives information from the server and closes the connection.
- The primary task of web server is to monitor a communication port on host machine, accept HTTP commands through that port and perform the operations specified by the commands.
- When the URL is received, it is translated into either a filename or a program name.

1.4.2 General Server Characteristics:

- The file structure of a web server has two separate directories
- The root of one of these is called **document root** which stores web documents
- The root of the other directory is called the **server root** which stores server and its support softwares
- The files stored directly in the document root are those available to clients through top level URLs
- The secondary areas from which documents can be served are called **virtual document trees.**
- Many servers can support more than one site on a computer, potentially reducing the cost of each site and making their maintenance more convenient. Such secondary hosts are called **virtual hosts.**
- Some servers can serve documents that are in the document root of other machines on the web; in this case they are called as **proxy servers**

1.4.3 Apache

- Apache is the most widely used Web server.
- The primary reasons are as follows: Apache is an excellent server because it is both fast and reliable.

- Furthermore, it is open-source software, which means that it is free and is managed by a large team of volunteers, a process that efficiently and effectively maintains the system.
- Finally, it is one of the best available servers for Unix-based systems, which are the most popular for Web servers.
- Apache is capable of providing a long list of services beyond the basic process of serving documents to clients.
- When Apache begins execution, it reads its configuration information from a file and sets its parameters to operate accordingly.

1.4.4 IIS

- Microsoft IIS server is supplied as part of Windows—and because it is a reasonably good server—most Windows-based Web servers use IIS.
- With IIS, server behaviour is modified by changes made through a window-based management program, named the IIS snap-in, which controls both IIS and ftp.
- This program allows the site manager to set parameters for the server.
- Under Windows XP and Vista, the IIS snap-in is accessed by going to *Control Panel, Administrative Tools, and IIS Admin.*

1.5 Uniform Resource Locators

- Uniform Resource Locators (URLs) are used to identify different kinds of resources on Internet.
- If the web browser wants some document from web server, just giving domain name is not sufficient because domain name can only be used for locating the server.
- It does not have information about which document client needs. Therefore, URL should be provided.
- The general format of URL is: **scheme: object-address**
- Example: **http://www.vtu.ac.in/results.php**
- The scheme indicates protocols being used. (http, ftp, telnet...)
- In case of http, the full form of the object address of a URL is as follows:
- **//fully-qualified-domain-name/path-to-document**
- URLs can never have embedded spaces
- It cannot use special characters like semicolons, ampersands and colons
- The path to the document for http protocol is a sequence of directory names and a filename, all separated by whatever special character the OS uses. (forward or backward slashes)
- The path in a URL can differ from a path to a file because a URL need not include all directories on the path
- A path that includes all directories along the way is called a **complete path.**

- Example: http://www.atme.in/
- In most cases, the path to the document is relative to some base path that is specified in the configuration files of the server. Such paths are called **partial paths.**
- Example: http://www.atme.in/

1.6 Multipurpose Internet Mail Extensions

- MIME stands for Multipurpose Internet Mail Extension.
- The server system apart from sending the requested document, it will also send MIME information.
- The MIME information is used by web browser for rendering the document properly.
- The format of MIME is: type/subtype
- Example: text/html , text/doc , image/jpeg , video/mpeg
- When the type is either text or image, the browser renders the document without any problem
- However, if the type is video or audio, it cannot render the document
- It has to take the help of other software like media player, win amp etc.,
- These software's are called as helper applications or plugins
- These non-textual information are known as HYPER MEDIA
- Experimental document types are used when user wants to create a customized information & make it available in the internet
- The format of experimental document type is: type/x-subtype
- Example: database/x-xbase , video/x-msvideo
- Along with creating customized information, the user should also create helper applications.
- This helper application will be used for rendering the document by browser.
- The list of MIME specifications is stored in configuration file of web server.

1.7 The Hyper Text Transfer Protocol

1.7.1 The Request Phase

The general form of an HTTP request is as follows:

1. HTTP method Domain part of the URL HTTP version
2. Header fields
3. Blank line
4. Message body

The following is an example of the first line of an HTTP request: **GET /storefront.html HTTP/1.1**

Table 1.1 HTTP request methods

Method	Description
GET	Returns the contents of the specified document
HEAD	Returns the header information for the specified document
POST	Executes the specified document, using the enclosed data
PUT	Replaces the specified document with the enclosed data
DELETE	Deletes the specified document

The format of a header field is the field name followed by a colon and the value of the field. There are four categories of header fields:

1. ***General***: For general information, such as the date
2. ***Request***: Included in request headers
3. ***Response***: For response headers
4. ***Entity***: Used in both request and response headers

A wildcard character, the asterisk (*), can be used to specify that part of a MIME type can be anything.

Accept: text/plain

Accept: text/html → *Can be written as* → Accept: text/*

The Host: *host name* request field gives the name of the host. The Host field is required for HTTP 1.1. The If-Modified-Since: *date* request field specifies that the requested file should be sent only if it has been modified since the given date. If the request has a body, the length of that body must be given with a Content-length field. The header of a request must be followed by a blank line, which is used to separate the header from the body of the request.

1.7.2 The Response Phase:

The general form of an HTTP response is as follows:

1. Status line
2. Response header fields
3. Blank line
4. Response body

The status line includes the HTTP version used, a three-digit status code for the response, and a short textual explanation of the status code. For example, most responses begin with the following:

HTTP/1.1 200 OK

The status codes begin with 1, 2, 3, 4, or 5. The general meanings of the five categories specified by these first digits are shown in Table 1.2.

Table 1.2 First digits of HTTP status codes

First Digit	Category
1	Informational
2	Success
3	Redirection
4	Client error
5	Server error

One of the more common status codes is one user never want to see: 404 Not Found, which means the requested file could not be found.

1.8 Security

Security is one of the major concerns in the Internet. The server system can be accessed easily with basic hardware support, internet connection & web browser. The client can retrieve very important information from the server. Similarly, the server system can introduce virus on the client system. These viruses can destroy the hardware and software in client. While programming the web, following requirements should be considered:

- **Privacy:** it means message should be readable only to communicating parties and not to intruder.
- **Integrity:** it means message should not be modified during transmission.
- **Authentication:** it means communicating parties must be able to know each other's identity
- **Non-repudiation:** it means that it should be possible to prove that message was sent and received properly

Security can be provided using cryptographic algorithm. Ex: private key, public key Protection against viruses and worms is provided by antivirus software, which must be updated frequently so that it can detect and protect against the continuous stream of new viruses and worms.

1.9 The Web Programmer's Toolbox

- Web programmers use several languages to create the documents that servers can provide to browsers.
- The most basic of these is **XHTML**, the standard mark-up language for describing how Web documents should be presented by browsers. Tools that can be used without specific knowledge of XHTML are available to create XHTML documents.
- A **plug-in** is a program that can be integrated with a word processor to make it possible to use the word processor to create XHTML. A **filter** converts a document written in some other format to XHTML.

- **XML** is a meta-mark-up language that provides a standard way to define new mark-up languages.
- **JavaScript** is a client-side scripting language that can be embedded in XHTML to describe simple computations. JavaScript code is interpreted by the browser on the client machine; it provides access to the elements of an XHTML document, as well as the ability to change those elements dynamically.
- **Flash** is a framework for building animation into XHTML documents. A browser must have a Flash player plug-in to be able to display the movies created with the Flash framework.
- **Ajax** is an approach to building Web applications in which partial document requests are handled asynchronously. Ajax can significantly increase the speed of user interactions, so it is most useful for building systems that have frequent interactions.
- **PHP** is the server-side equivalent of JavaScript. It is an interpreted language whose code is embedded in XHTML documents. PHP is used primarily for form processing and database access from browsers.
- **Servlets** are server-side Java programs that are used for form processing, database access, or building dynamic documents. JSP documents, which are translated into servlets, are an alternative approach to building these applications. JSF is a development framework for specifying forms and their processing in JSP documents.
- **ASP.NET** is a Web development framework. The code used in ASP.NET documents, which is executed on the server, can be written in any .NET programming language.
- **Ruby** is a relatively recent object-oriented scripting language that is introduced here primarily because of its use in Rails, a Web applications framework.
- **Rails** provides a significant part of the code required to build Web applications that access databases, allowing the developer to spend his or her time on the specifics of the application without the drudgery of dealing with all of the housekeeping details.

1.10 Origins and Evolution of HTML and XHTML

HTML → Hyper Text Mark-up Language

XHTML → eXtensible Hyper Text Mark-up Language

1.10.1 HTML versus XHTML

HTML	XHTML
HTML is much easier to write	XHTML requires a level of discipline many of

	us naturally resist
huge number of HTML documents available on the Web, browsers will continue to support HTML as far as one can see into the future.	some older browsers have problems with some parts of XHTML.
HTML has few syntactic rules, and HTML processors (e.g., browsers) do not enforce the rules it does have. Therefore, HTML authors have a high degree of freedom to use their own syntactic preferences to create documents. Because of this freedom, HTML documents lack consistency, both in low-level syntax and in overall structure.	XHTML has strict syntactic rules that impose a consistent structure on all XHTML documents. Another significant reason for using XHTML is that when you create an XHTML document, its syntactic correctness can be checked, either by an XML browser or by a validation tool
Used for displaying the data	Used for describing the data

1.10.2 Basic Syntax

- The fundamental syntactic units of HTML are called ***tags***.
- In general, tags are used to specify categories of content.
- The syntax of a tag is the tag's name surrounded by *angle brackets* (< and >).
- Tag names must be written in all lowercase letters.
- Most tags appear in pairs: an *opening tag and a closing tag*.
- The name of a closing tag is the name of its corresponding opening tag with a slash attached to the beginning. For example, if the tag's name is p, the corresponding closing tag is named /p.
- Whatever appears between a tag and its closing tag is the ***content*** of the tag. Not all tags can have content.
- The opening tag and its closing tag together specify a container for the content they enclose.
- The container and its content together are called an ***element***.
- Example: ** This is ATME Web Programming Notes. **
- The paragraph tag, , marks the beginning of the content; the tag marks the end of the content of the paragraph element.
- Attributes, which are used to specify alternative meanings of a tag, can appear between an opening tag's name and its right angle bracket.
- They are specified in keyword form, which means that the attribute's name is followed by an equal's sign and the attribute's value.
- Attribute names, like tag names, are written in lowercase letters.
- Attribute values must be delimited by double quotes.

- Comments in programs increase the readability of those programs. Comments in XHTML have the same purpose. They can appear in XHTML in the following form:
- **<!- - anything except two adjacent dashes - ->**
- Browsers ignore XHTML comments—they are for people only. Comments can be spread over as many lines as are needed. For example, you could have the following comment:
- **<!- - CopyRights.html**

  This notes is prepared by Kswamy of Computer Science Department ATME, Mysore - ->

1.10.3 Standard XHTML Document Structure

- Every XHTML document must begin with an xml declaration element that simply identifies the document as being one based on XML. This element includes an attribute that specifies the version number 1.0.
- The xml declaration usually includes a second attribute, encoding, which specifies the encoding used for the document [utf-8].
- Following is the xml declaration element, which should be the first line of every XHTML document:

<?xml version = "1.0" encoding = "utf-8"?>

- Note that this declaration must begin in the first character position of the document file.
- The xml declaration element is followed immediately by an SGML DOCTYPE command, which specifies the particular SGML document-type definition (DTD) with which the document complies, among other things.
- The following command states that the document in which it is included complies with the XHTML 1.0 Strict standard:

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml1-strict.dtd">

- An XHTML document must include the four tags <html>, <head>, <title>, and <body>.
- The <html> tag identifies the root element of the document. So, XHTML documents always have an <html> tag immediately following the DOCTYPE command, and they always end with the closing html tag, </html>.
- The html element includes an attribute, xmlns, that specifies the XHTML namespace, as shown in the following element:

<html xmlns = "http://www.w3.org/1999/xhtml">

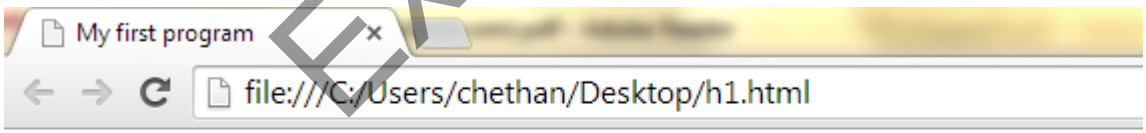
- Although the xmlns attribute's value looks like a URL, it does not specify a document. It is just a name that happens to have the form of a URL.
- An XHTML document consists of two parts, named the *head* and the *body*.
- The <head> element contains the head part of the document, which provides information about the document and does not provide the content of the document.

- The body of a document provides the content of the document.
- The content of the title element is displayed by the browser at the top of its display window, usually in the browser window's title bar.

1.11 Basic Text Markup

We will have a look at a complete XHTML document:

```
<?xml version = "1.0" encoding = "utf-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml1-strict.dtd">
<!- - complete.html
A document which must be followed throughout the notes - ->
<html xmlns = "http://www.w3.org/1999/xhtml">
<head>
<title>
My first program
</title>
</head>
<body>
<p>
My Dear ATME Friends, All The Best..!! Have a Happy Reading of my notes..!!
</p>
</body>
</html>
```



My Dear ATME Friends, All The Best..!! Have a Happy Reading of my notes..!!

PLEASE NOTE: From here onwards programming in XHTML will begin. Please add the following compulsory document structure to all programs in the first 4 lines and skip the simple <html> tag of first line because I have begun the coding part directly .

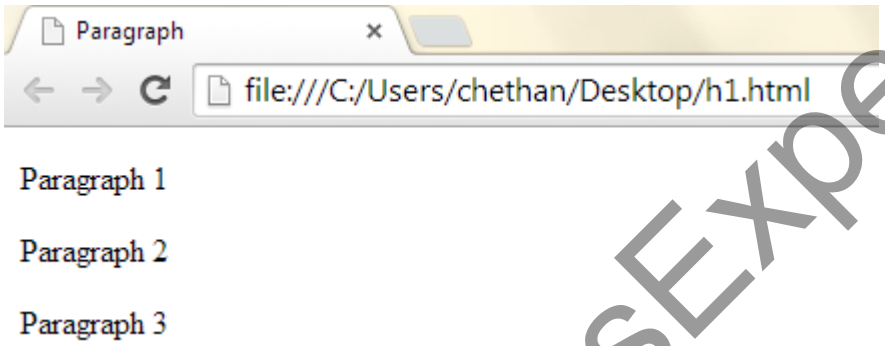
<?xml version = "1.0" encoding = "utf-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml1-strict.dtd">

```
<html xmlns = "http://www.w3.org/1999/xhtml">
```

1.11.1 Paragraphs:

It begins with and ends with . Multiple paragraphs may appear in a single document.

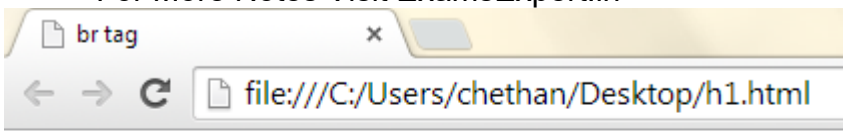
```
<html>
 <head>
        <title> Paragraph </title>
</head>
<body>
         <p> Paragraph 1 </p>
        <p> Paragraph 2 </p>
        <p> Paragraph 3 </p>
</body>
 </html>
```



1.11.2 Line Breaks:

The break tag is specified as
. The slash indicates that the tag is both an opening and closing tag.

```
<html>
<head>
         <title> br tag </title>
</head>
<body>
        <p> My Name is Chethan <br/>
        I am from CSE Department <br/>
        ATME, Mysore </p>
 </body>
 </html>
```



My Name is Chethan
I am from CSE Department
ATME, Mysore

1.11.3 Preserving White Space:

Sometimes it is desirable to preserve the white space in text—that is, to prevent the browser from eliminating multiple spaces and ignoring embedded line breaks. This can be specified with the **<pre>** tag.

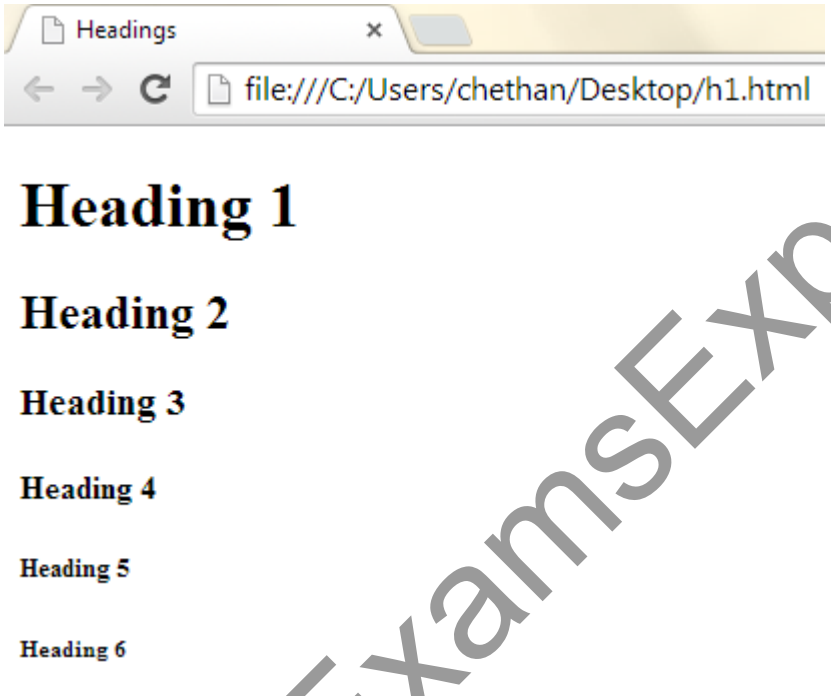
```
<html>
 <head>
        <title> Pre Tag </title>
 </head>
<body>
        <p><pre> My Name is Chethan
                                I am from CSE Department
                                        ATME, Mysore </pre></p>
</body>
</html>
```

1.11.4 Headings:

- In XHTML, there are six levels of headings, specified by the tags <h1>, <h2>, <h3>, <h4>, <h5>, and <h6>, where <h1> specifies the highest-level heading.
- Headings are usually displayed in a boldface font whose default size depends on the number in the heading tag.
- On most browsers, <h1>, <h2>, and <h3> use font sizes that are larger than that of the default size of text, <h4> uses the default size, and <h5> and <h6> use smaller sizes.
- The heading tags always break the current line, so their content always appears on a new line.
- Browsers usually insert some vertical space before and after all headings.

```
<html>
```

```
<head>
    <title> Headings </title>
</head>
<body>
     <h1> Heading 1 </h1>
    <h2> Heading 2 </h2>
    <h3> Heading 3 </h3>
    <h4> Heading 4 </h4>
     <h5> Heading 5 </h5>
     <h6> Heading 6 </h6>
 </body>
 </html>
```

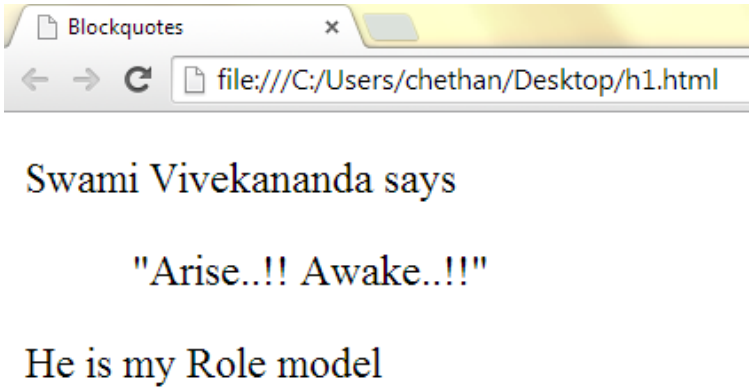


1.11.5 Block Quotations:

The <blockquote> tag is used to make the contents look different from the surrounding text.

```
<html>
<head>
    <title> Blockquotes </title>
</head>
<body>
        <p> Swami Vivekananda says </p>
     <blockquote>
        <p> "Arise..!! Awake..!!" </p>
    </blockquote>
```

```
        <p> He is my Role model </p>
</body>
</html>
```



1.11.6 Font Styles and Sizes:

- <b>, <i> and <u> specifies bold, italics and underline respectively.
- The emphasis tag, , specifies that its textual content is special and should be displayed in some way that indicates this distinctiveness. Most browsers use italics for such content.
- The strong tag, is like the emphasis tag, but more so. Browsers often set the content of strong elements in bold.
- The code tag, <code>, is used to specify a monospace font, usually for program code.

```
<html>
<head>
 <title> font styles and sizes </title>
</head>
<body>
        <p>
                <pre> Illustration of Font Styles
                        <b> This is Bold </b>
                        <i> This is Italics </i>
                        <u> This is Underline </u>
                        <em> This is Emphasis </em>
                        <strong> This is strong </strong>
                        <code> Total = Internals + Externals //this is code</code>
                </pre>
        </p>
         <p>
                <pre> Illustration of Font Sizes (subscripts and superscripts)
                        x<sub>2</sub><sup>3</sup> + y<sub>1</sub><sup>2</sup>
```

```
        </pre>
    </p>
</body>
</html>
```

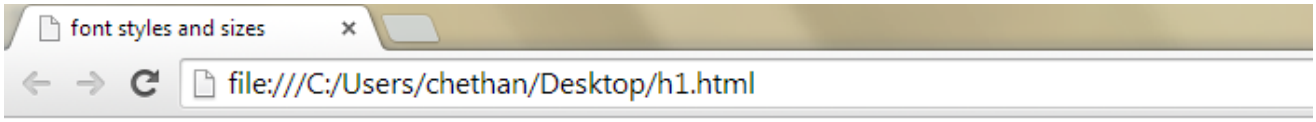


Illustration of Font Styles
This is Bold
This is Italics
This is Underline
This is Emphasis
This is strong
Total = Internals + Externals //this is code

Illustration of Font Sizes (subscripts and superscripts)
$x_2^3 + y_1^2$

1.11.7 Character Entities:

- XHTML provides a collection of special characters that are sometimes needed in a document but cannot be typed as themselves.
- In some cases, these characters are used in XHTML in some special way—for example, >, <, and &. In other cases, the characters do not appear on keyboards, such as the small raised circle that represents "degrees" in a reference to temperature.
- These special characters are defined as entities, which are codes for the characters. An entity in a document is replaced by its associated character by the browser.

```
<html>
 <head>
        <title> Character Entities </title>
 </head>
<body>
        <p>
                <pre> Illustration of character entities
                        if you get &gt 70%, then you will get FCD
                        if you get &lt 35%, then you are Fail
                        &frac12 of my classmates get very good marks
                        Now, the temperature in Bangalore is 30&deg C
                </pre>
```

```
    </p>
</body>
</html>
```

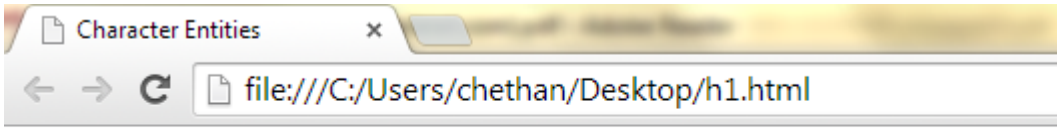
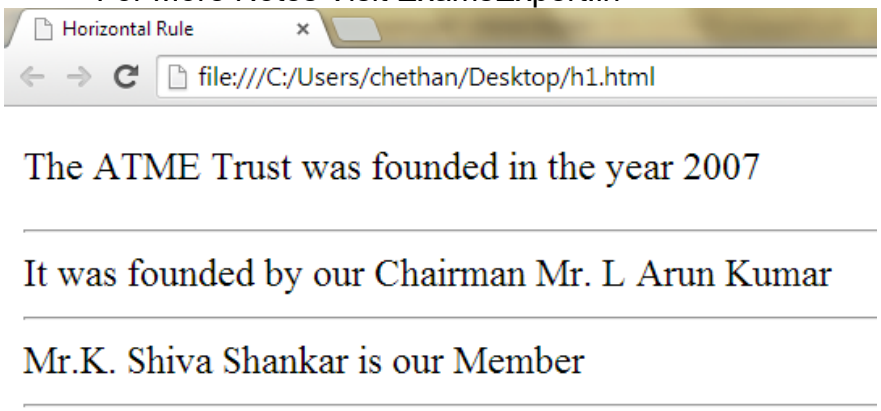


Illustration of character entities
if you get > 70%, then you will get FCD
if you get < 35%, then you are Fail
½ of my classmates get very good marks
Now, the temperature in Bangalore is 30° C

1.11.8 Horizontal Rules:

- The parts of a document can be separated from each other, making the document easier to read, by placing horizontal lines between them. Such lines are called horizontal rules.
- The block tag that creates them is <hr />. The <hr /> tag causes a line break (ending the current line) and places a line across the screen. Note again the slash in the <hr /> tag, indicating that this tag has no content and no closing tag.

```
<html>
 <head>
        <title> Horizontal Rule </title>
 </head>
 <body>
        <p>
              The ATME Trust  was founded in the year 2007 <hr/>
              It was founded by our Chairman Mr. L Arun Kumar <hr/>
              Mr.K. Shiva Shankar  is our Member <hr/>
        </p>
</body>
</html>
```



1.11.9 The meta Element:

- The meta element is used to provide additional information about a document. The meta tag has no content; rather, all of the information provided is specified with attributes.
- The two attributes that are used to provide information are name and content. The user makes up a name as the value of the name attribute and specifies information through the content attribute.
- One commonly chosen name is keywords; the value of the content attribute associated with the keywords are those which the author of a document believes characterizes his or her document.
- An example is
  - **<meta name = "Title" content = "Programming the Web" />**
  - **<meta name = "Author" content = "Divya K" />**
- Web search engines use the information provided with the meta element to categorize Web documents in their indices.

1.12 IMAGES

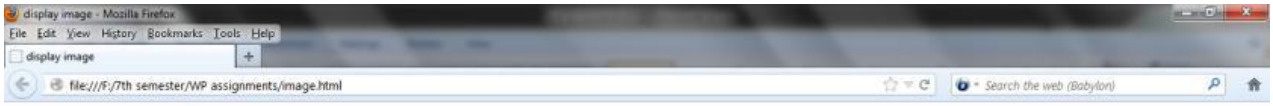
- Image can be displayed on the web page using <img> tag.
- When the <img> tag is used, it should also be mentioned which image needs to be displayed. This is done using src attribute.
- Attribute means extra information given to the browser
- Whenever <img> tag is used, alt attribute is also used.
- Alt stands for alert.
- Some very old browsers would not be having the capacity to display the images.
- In this case, whatever is the message given to alt attribute, that would be displayed.
- Another use of alt is → when image display option has been disabled by user. The option is normally disabled when the size of the image is huge and takes time for downloading.

<html>

<head>

<title>display image</title>

```
</head>
    <body>
        <img src="java.png" alt="cannot display"/>
    </body>
</html>
```



NOTE:

JPEG → Joint Photographic Experts Group

GIF → Graphic Interchange Format

PNG → Portable Network Graphics

XHTML Document Validation:

The W3C provides a convenient Web-based way to validate XHTML documents against its standards.

Step 1: The URL of the service is http://validator.w3.org/file-upload.html. Copy & paste this link.

Step 2: You will be driven to "Validate by File Upload" option automatically.

Step 3: Browse for a XHTML program file in your computer. (example: *F:/complete.html*)

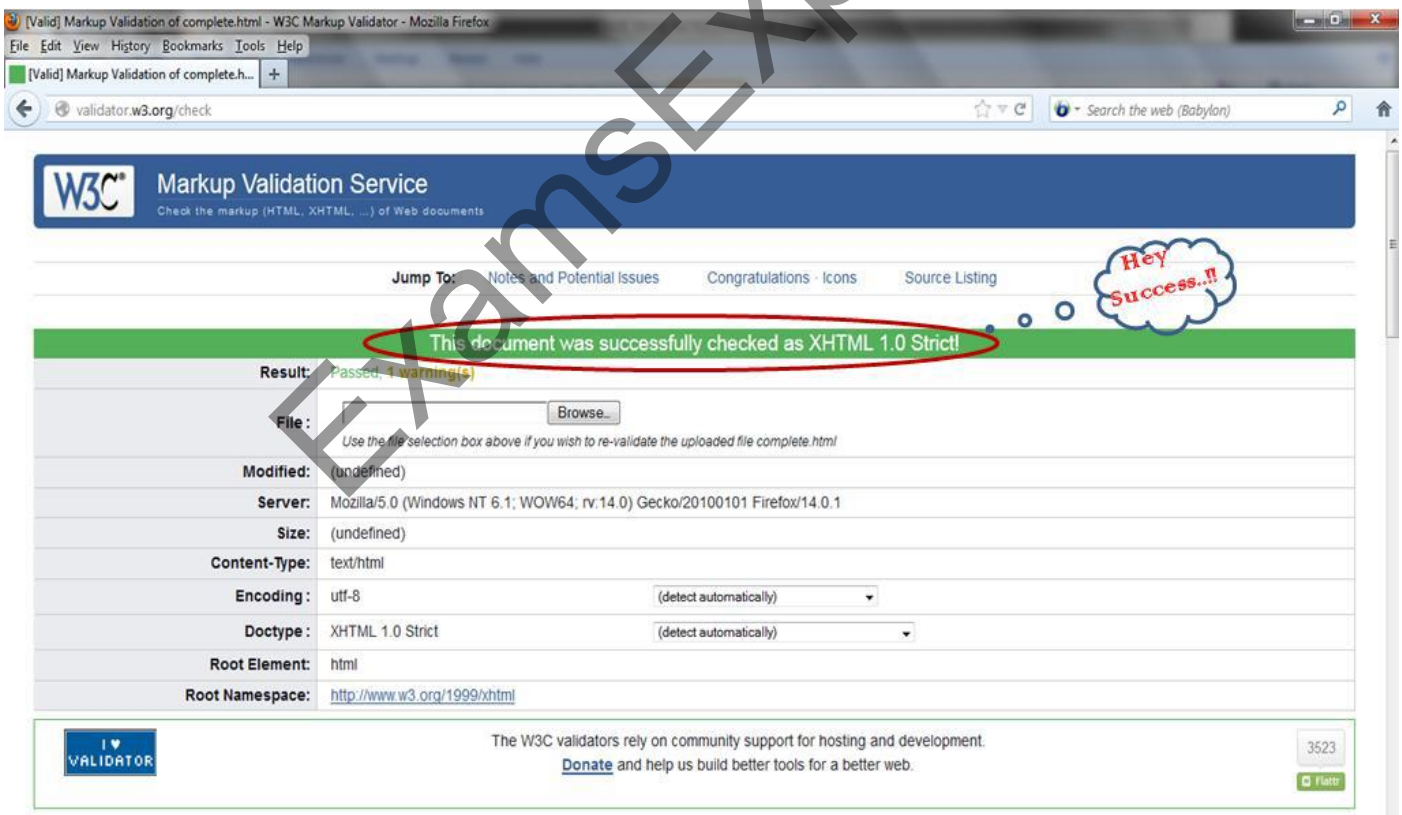
Step 4: Click on "More Options" and select your criteria like *show source*

Step 5: After all the settings, click on "Check" button

Now you will be navigated to another page which shows success or failure.

In our example, the file *complete.html* is a valid XHTML file. So the output shows success..!!

OUT PUT:



1.13 HYPERTEXT LINKS

1.13.1 Links:

- Hyperlinks are the mechanism which allows the navigation from one page to another.
- The term “hyper” means beyond and “link” means connection
- Whichever text helps in navigation is called hypertext
- Hyperlinks cam be created using **<a>** (anchor tag)
- The attribute that should be used for <a> is **href**

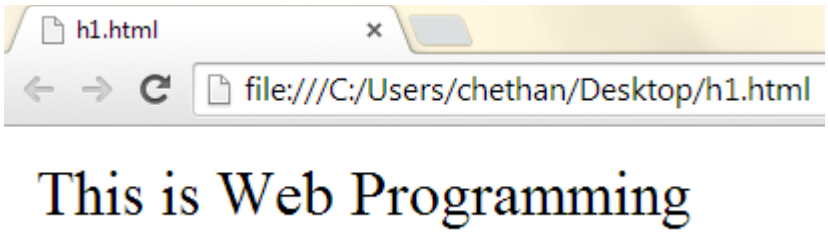
Program: *hyper.html*

```
<html>
<head>
<title> hyperlink </title>
</head>
<a href = "link.html"> CLICK HERE
</a>
</html>
```

Program: *link.html*

```
<html>
<body> This is Web Programming
</body>
</html>
```

After clicking on the above text, we can navigate to another page “link.html” as shown below



1.13.2 Targets within Documents:

If the target of a link is not at the beginning of a document, it must be some element within the document, in which case there must be some means of specifying it.

The target element can include an id attribute, which can then be used to identify it in an href attribute. (observe the scroll bar in the outputs given)

```
<html>
<head>
    <title> target link</title>
</head>
 <body>
    <h2 id = “avionics”> Avionics </h2>
    <a href = “#avionics”> What about avionics? </a>
    <a href = “AIDAN1.html#avionics”> Avionics </a>
</body>
</html>
```

Question paper question

1) Give syntax and an example for each of the following tags. 1.<pre> 2.<a> 3.<img> 4.<sub> 5.<p> (10 M)
2) Explain with an example the following tags 1.select 2.frames 3.colspan 4.radio button 5.style class selector (10 M)
3) Give and explain syntax of following tags 1.<blockquote> 2.meta (03 M)
4) Explain the following tags with examples i. <img> ii. <a> (04 M)
5) Syntax and an example for each of the following tags. 1.<pre> 2.<p> 3.<sup> 4.<sub> 5.<blackquote> (10 m)
6) Give the standard structure of XHTML document. How line breaks, heading and fonts are handled in XHTML? (10 M)
7) Explain standard XHTML document structure (08 M)
8) Explain the different image formats, write XHTML document to illustrate use of <img> (with all its attributes) (08 M)
9) Give the syntactic difference between HTML and XHTML (08 M)
10) Discuss the following tags with syntax and example i) <pre> ii. <meta> (04 M)