2.2 HTML – AN INTRODUCTION

The backbone of the World Wide Web is made of *HTML files,* which are specially-formatted documents that can contain links, as well as images and other media.

HTML stands for HyperText Markup Language, and it is the most widely used language to design Web Pages.

A webpage is accessed by entering a URL *(Uniform Resource Locator- a mechanism used to locate a webpage on the internet.)* addresses in the *Address field* of web browser. A webpage may contain text, graphics, and hyperlinks to other web pages and files.

A *website* is composed of a group of web pages linked together. It is a central location that contains more than one web page.

HTML (Hypertext Markup Language), as its name suggests, is a markup language.

- ***Hypertext*** refers to the way in which Web pages (HTML documents) are linked together. When you click a link in a Web page, you are using hypertext. So, hypertext is simply a piece of text that works as a link.
- ***Markup Language*** describes how HTML works. With a markup language, you simply "mark up" a text document with tags that tell a Web browser how to structure it to display. So, markup language is a way of writing layout information within documents.

Note: HTML documents are also called *web pages*.

2.2.1 BRIEF HISTORY OF HTML

HTML was created by *Berners-Lee* in late 1991 but "HTML 2.0" was the first standard HTML specification which was published in 1995.

HTML 4.01 was a major version of HTML and it was published in late 1999.

Though HTML 4.01 version is widely used but currently we are having HTML-5 version which is an extension to HTML 4.01, and this version was published in 2012.

2.2 BASIC HTML CONCEPTS

In HTML, markup commands *(tags)* applied to your web-based content tell the web browser the structure of the document i.e. how and when the content to be displayed.

For example, if you want to display a section of text in boldface, you surround the corresponding text with the boldface markup tags, <B> and </B>, as shown below:

```
<B>Welcome you all</B>
```

When the browser reads a document that has HTML markup in it, it determines how to render it onscreen by considering the HTML elements embedded within the document.

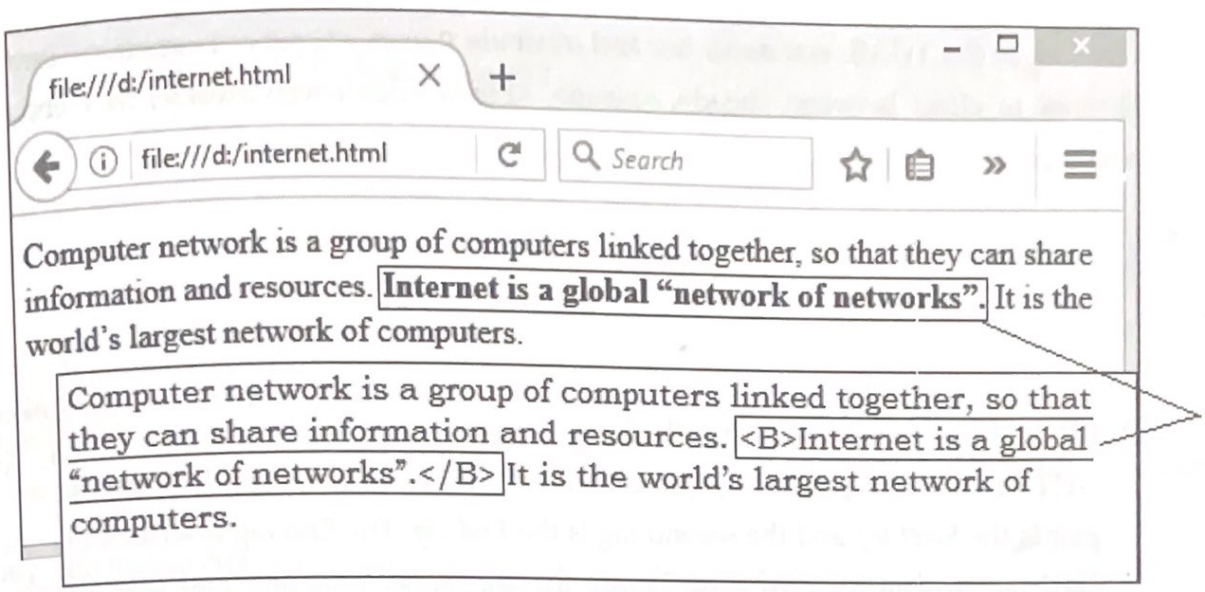


Fig. 2.1 (Interpretation of a Web page with HTML Markup)

An HTML document is simply a text file that contains the information you want to publish over the web. It also contains embedded instructions, called *elements* that indicate how a Web browser should structure or present the document.

2.4 AN OVERVIEW OF HTML MARKUP

A markup language such as HTML is simply a collection of *codes* or *elements* that are used to indicate the structure and format of a document. The codes have meaning that is interpreted by a Web browser, which renders the document. Elements in HTML document consist of alphanumeric tags within angle brackets.

2.4.1 HTML TAGS

A webpage can contain text, graphics, audio, video and animations. We can mark up such contents to indicate how they should appear over a web page. To mark up the contents, we can use various commands, called *Markup Indicators* or *TAGS*.

So, HTML Tags are basically HTML based commands or keywords surrounded by angle brackets like <HTML>.

Every tag consists of a tag name, sometimes followed by an optional list of tag *attributes* or *properties*, all placed between opening and closing brackets (< and >). The simplest tag is nothing more than a name appropriately enclosed in brackets, such as <HEAD> and <I>. More complicated tags contain one or more attributes, which specify or modify the behavior of the tag.

According to the HTML standard, tag and attribute names are *not case-sensitive*. There's no difference in effect between <head>, <Head>, <HEAD>, or even <HeaD>; they are all equivalent.

2.4.1.1 Types of Tags

(1) Paired / Container Tags

(2) Unpaired / Empty Tags

(1) Paired / Container Tags: HTML Tags normally come in pairs like <HTML> and </HTML>. Such types of tags are called *Paired or Container Tags*. The first tag in a pair is the *Start tag* and the second tag is the *End tag*. The End tag is written like the Start tag, with a forward slash before the tag name. Start and End tags are also called *Opening tags* and *Closing tags*.

Format of Paired / Container Tag:

```
<tagname>content</tagname>
```

Example:

```
<B>HELLO</B>
```

(2) Unpaired / Empty Tags: An unpaired or empty tag refers to a tag, which does not have an ending/closing tag.

Format of Unpaired / Empty Tag:

```
<tagname>
```

Example:

```
<BR>
```

Note: HTML tags are not case-sensitive.

2.4.2 HTML ELEMENTS

An HTML element is everything from the *start tag* to the *end tag*, including the *contents*.

Example:

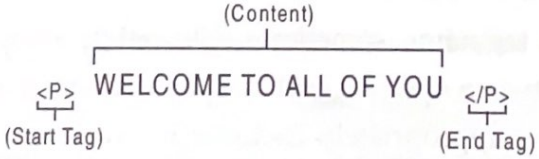


Fig. 2.2 (HTML Element)

2.4.2.1 Nested HTML Elements

Most HTML elements can be nested i.e. they can contain other HTML elements.

Example:

```
<HTML>
<BODY>
 <B> WELCOME TO ALL OF YOU. </B>
</BODY>
</HTML>
```

In this example,

- The <BODY> element defines the body of the HTML document.
- The element has a start tag <BODY> and an end tag </BODY>.
- The content of <BODY> element is another HTML element (<B> element, defines the bold effect over the text.).

2.4.2.2 Empty HTML Elements

HTML elements with no content are called *Empty elements.*

Example:

 is an empty element without a closing tag (the
 tag defines a line break).

Note: Adding a slash inside the start tag, like
, is the proper way of closing empty elements.

2.4.3 HTML ATTRIBUTES

Attributes are important part of HTML Elements. An attribute is used to define the characteristics of an element and is always placed inside the element's opening tag. Attributes provide some sort of additional information about an element.

All attributes are made up of two parts: a *name* and a *value*. Where, the name is the property you want to set and the value is what you want the value of the property to be. The value of the attribute should be put in double quotation marks, and is separated from the name by the equals sign.

In simple terms, attributes come in name/value pairs like: name="value".

Example:

```
<BODY BGCOLOR="RED">
```

In this example, <BODY> element carries an attribute whose name is BGCOLOR, which you can use to set the background color of the web page. Here, the value set to BGCOLOR attribute is "RED". This means the background color of the web page will be RED.

Note: Attribute values should always be enclosed in quotes.

2.5 STRUCTURE OF HTML DOCUMENT

Every HTML document or page has a basic structure. An HTML document starts and ends with <HTML> and </HTML> tags. These tags tell the browser that the entire document is composed in HTML. Inside these two tags, the document is split into two sections:

(i) Header Section

(ii) Body Section

(i) Header Section: This section define the <HEAD>...</HEAD> elements, which contain information about the document such as title of the document, author of the document etc. Information inside this tag does not display outside.

(ii) Body Section: This section define the <BODY>...</BODY> elements, which contain the real content of the document that you see on your screen.

Below is a visualization of an HTML page structure:

Fig. 2.3 (Basic Structure of HTML Document)

2.6 CREATING HTML DOCUMENT

Creating an HTML document is a very easy process. To create an HTML document, generally you need the following software tools:

(i) Text Editor: To write or edit the HTML code, you need a simple text-editor, like-

Notepad (A standard program comes with Windows operating system.). You can also use some professional HTML editors like- *Adobe Dreamweaver, Microsoft FrontPage* etc.

(ii) **Web Browser:** The purpose of a web browser (such as *Google Chrome, Internet Explorer, Firefox, Safari*) is to read HTML documents and display them as web pages. The browser does not display the HTML tags, but uses the tags to determine how the content of the HTML page is to be presented or displayed to the user.

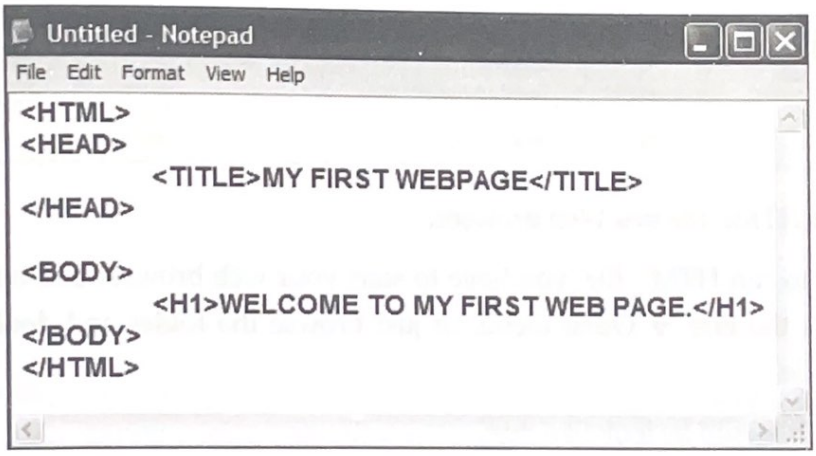
Here are the simple steps to create a basic HTML document (Using Notepad):

Step 1: Open **Notepad.**

- To open Notepad, Click **Start →All Programs → Accessories → Notepad.**

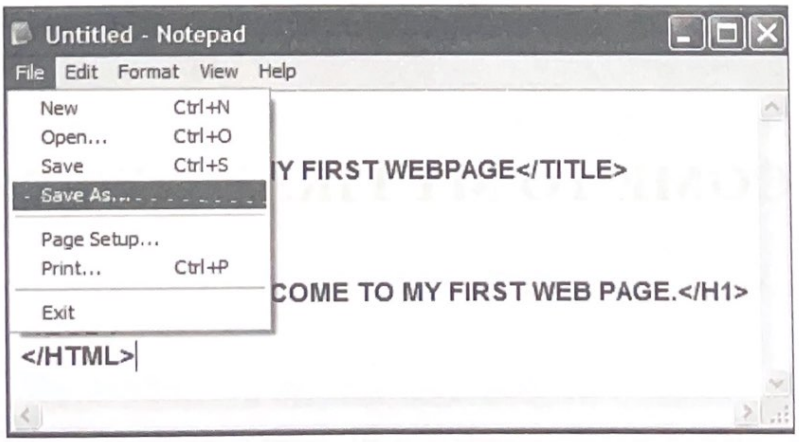
Step 2: Write the **HTML Code.**

- Type your HTML code into Notepad, like:

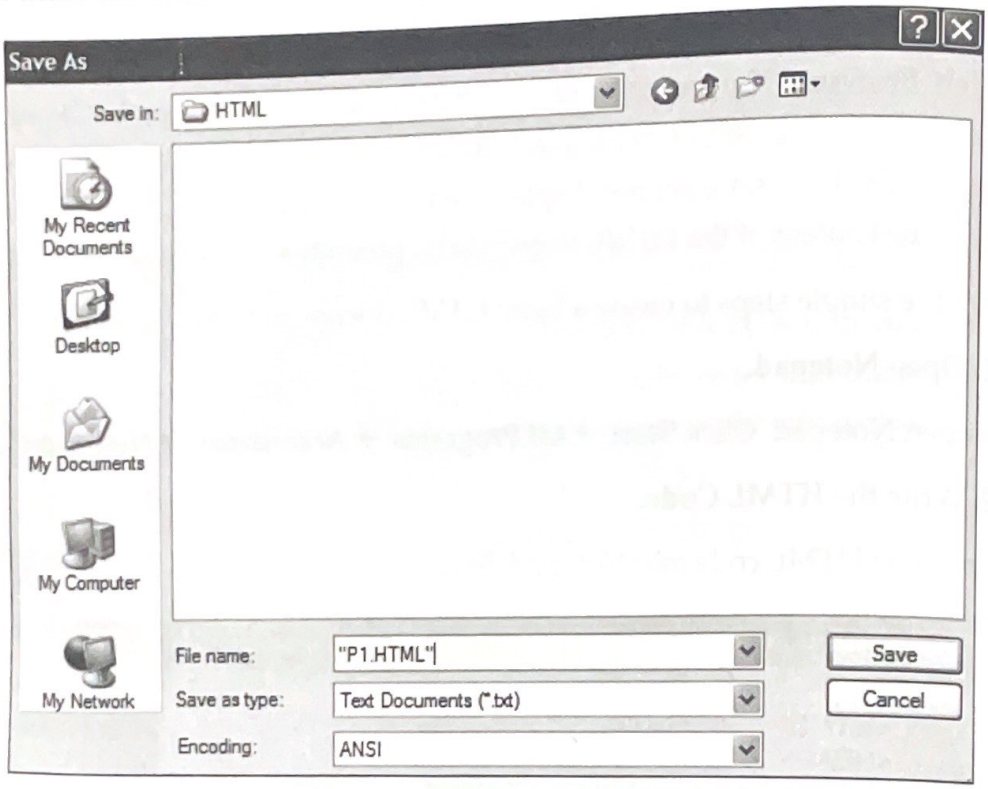


Step 3: Save HTML Code.

- To save HTML code, Select **Save as...** in Notepad's File menu.

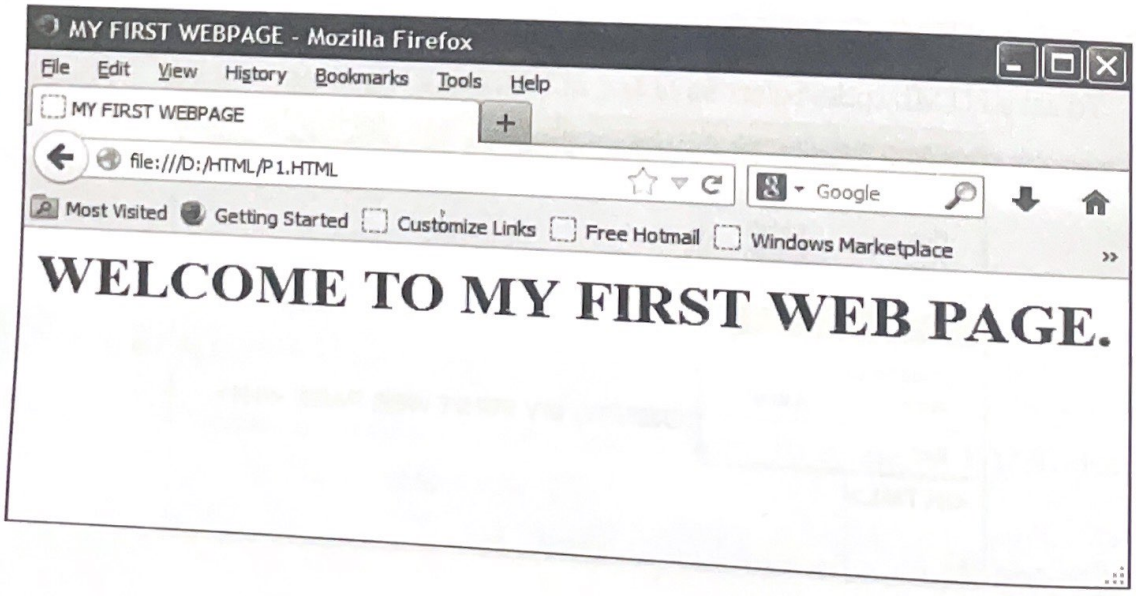


- **Save As dialog box** will be appeared. Save your HTML file with either the **.htm** or the **.html** file extension under an appropriate folder.



Step 4: Run HTML file in a Web Browser.

- Finally, to run HTML file, you have to start your web browser and open your html file from the **File → Open** menu, or just browse the folder and double-click your HTML file.
- The result would be look like this:



2.7 DOCUMENT TAGS

Document tags define the overall structure of an HTML document. There are four tags every HTML document should have. These tags define what type of document it is, and the major sections of the document. These tags are:

(i) <HTML>

(ii) <HEAD>

(iii) <TITLE>

(iv) <BODY>

(i) <HTML>

The <HTML> tag is the main container or containing element for the whole HTML document. It represents the root of an HTML document. It simply tells the browser that this is an HTML document. Each HTML document should have one <HTML> tag and each document should end with a closing </HTML> tag.

Following two elements appear as direct children of an <HTML> tag:

- <HEAD>
- <BODY>

(ii) <HEAD>

The <HEAD> tag is just a container for all other header elements. It simply defines the header section of the HTML document. It should be the first thing to appear after the opening <HTML> tag. Each <HEAD> tag should contain a <TITLE> tag indicating the title of the document. Although it may also contain any combination of the following tags, in any order:

(a) <LINK>: The <LINK> tag is used to link to an external file, such as a style sheet or JavaScript file.

(b) <STYLE>: The <STYLE> tag is used to include CSS rules inside the HTML document.

(c) <SCRIPT>: The <SCRIPT> tag is used to include JavaScript or VBScript inside the HTML document.

(d) <META>: The <META> tag includes information about the document such as keywords and a description, which are particularly helpful for search applications.

(iii) <TITLE>

The <TITLE> tag is used to set a title for every HTML document. The <TITLE> tag defines a title in the browser's title bar. This element is a child of the <HEAD> element. The <TITLE> tag should contain only the text for the title and it may not contain any other elements.

(iv) <BODY>

The <BODY> tag appears after the <HEAD> tag. It defines the body section of HTML document i.e. it simply defines the document's body. The <BODY> tag contains all the contents of an HTML document, such as text, hyperlinks, images, tables, lists, etc.

2.7.1 <BODY> Attributes

Attribute	Value	Description
alink	color_name	Specifies the color of an active link in a document.
background	URL	Specifies a background image for a document.
bgcolor	color_name	Specifies the background color of a document.
link	color_name	Specifies the color of unvisited links in a document.
text	color_name	Specifies the color of the text in a document.
vlink	color_name	Specifies the color of visited links in a document.

Example:

```
<HTML>
<HEAD>
        <TITLE>DOCUMENT TAGS</TITLE>
</HEAD>
<BODY BGCOLOR = "GREEN" TEXT = "BLUE">
        HAVE A NICE DAY.
</BODY>
</HTML>
```