

## INTRODUCTION

- HTML is an evolving language, and stands for Hyper Text Markup Language
- It is a text file containing small **markup tags** which are instructions given to browser about how to display the page.
  - An HTML file must have an **htm** or **html** file extension.
  - An HTML file can be created using a **simple text editor** (like Notepad in windows).

HTML was invented by Tim Berners-Lee while at CERN, the European Laboratory for Particle Physics in Geneva.

## 4.1 WHY HTML

Hypertext is an ordinary text with formatting facilities and Markup is the process of taking ordinary text and adding extra features to it.

Thus the HTML documents are text files made up of HTML elements that define a document and guide its display. In practical terms, HTML is a collection of *platform-independent styles* that define the various components of a World Wide Web document. HTML provides the user with a consistent interface and a highly effective medium for presenting information to developers.

HTML is a set of instructions given to Web browser for formatting and layout of WebPages, so it may be possible that a look of the same HTML code may differ since different browsers may interprets them differently.

## Advantages of HTML

- 1) HTML is an easy to use, learn, implement and flexible alternative to traditional presentation and tedious software.
- 2) Contains powerful formatting facilities.
- 3) HTML documents are device and platform independent. (Since it can be designed to work on not only home PCs but also on graphical workstations, dumb terminals, network computers, hand-held devices etc.)
- 4) You can traverse to any HTML document required because of hyper linking facility available, thus controlled navigation is possible.
- 5) Required HTML pages can be updated easily, without changing whole document.
- 6) It is a kind of software, which has been called world ware.
- 7) Independent work can be done and you need not rely on application or program vendor.
- 8) No expensive license software or hardware required.

- 9) If compatibility with user habits, expectations and multiple platforms is the goal, then HTML is the only approach to delivering a web application.

### **Disadvantages of HTML**

- 1) HTML doesn't offer programming languages features and capabilities.
- 2) It's easy to write "bad" HTML containing errors.
- 3) Complex HTML code is hard to read and understand and code complexity increases to make interactive web page. So building complex pages is very time consuming.
- 4) It's easy to make mistakes (e.g. leaving out a ">" or "/" character).
- 5) Special types of software like scripting languages (VB Script, Java Script) are required for handling different events and validations.
- 6) Can't detect errors easily since no special debugging tool is provided.

## **4.2 HTML DOCUMENTATION**

HTML documents are plain-text (also known as ASCII) files that can be created using any text editor (e.g., Emacs or vi on UNIX machines; Simple Text on a Macintosh; Notepad on a Windows machine). You can also use word-processing software if you remember to save your document as "text only with line breaks".

### **HTML Editors**

HTML Editors are programming tools for Hyper Text Markup Language (HTML) documents.

There are three categories of HTML Editors:

1. Text Editors
2. HTML Code Editors
3. HTML Design Tools

#### **1) Text Editors**

These editors only edit ASCII text. They offer no functionality to facilitate better HTML development. They are useful if your knowledge of HTML is excellent. Some examples of Text Editors include Notepad (Windows), Simple Text (Macintosh), and Pico (Unix).

They are typically WYSIWYG. WYSIWYG is an acronym for "what you see is what you get"; it means that you design your HTML document visually, as if you were using a word processor, instead of writing the markup tags in a plain-text file and imagining what the resulting page will look like.

#### **2) HTML Code Editors**

These editors may or may not be WYSIWYG.

### 3) HTML Design Tools

These tools are intended for HTML development without exposing the code to the author. They are typically WYSIWYG. Many of these tools do allow the user to access the HTML code, however this is not usually apparent to the user. Some examples of HTML Design Tools include NetObjects Fusion (Web Development), and Microsoft Office 97 (Traditional Office/Design Tools that provide HTML output).

You can concentrate on the content, rather than the syntax, of your Web site. You can Create a Web site without learning HTML. You can Design Elegant and Consistent Web sites with a few key strokes, since they are more user friendly

### ➤ HTML ELEMENT

An *element* is a fundamental component of the structure of a text document. Some examples of elements are heads, tables, paragraphs, and lists. Elements can contain plain text, other elements, or both.

HTML elements are defined using HTML tags.

### ➤ HTML Tags

- HTML tags are used to mark-up HTML elements.
- HTML tags are surrounded by the two characters < and >
- The surrounding characters are called angle brackets.
- HTML tags normally come in pairs like <b> and </b> they are usually paired to start and end the tag instruction.
- The first tag in a pair is the start tag, the second tag is the end tag.
- The text between the start and end tags is the element content.
- HTML tags are not case sensitive; <b> means the same as <B>.

Syntax is:

<tag name> text (element content) </tag name>  
<tag name attribute name="argument/value"> text (element content)  
</tag name>  
or just  
<tag name>

For example:

<TITLE> WELCOME TO MY WEBSITE </TITLE>

<A HREF="URL"> CLICK HERE </A>

An HTML document is composed of a single element:

<HTML> ... </HTML>

that is, in turn, composed of head and body elements:

<HEAD> ... </HEAD>

and <BODY> ... </BODY>

### ➤ Tag Attributes

Tags can have attributes. Attributes can provide additional information about the HTML elements on your page, included inside the start tag. Attributes always come in **name/value pairs** like this: name="value" Syntax is

<tag name attribute name="value">

e.g. <IMG SRC="c:\a.jpg">

so here src is an attribute for image tag. Normally its value is specified in quotes.

### • Some useful Tips while writing HTML document

HTML is **not case sensitive**. So <head> is equivalent to <HEAD> or <Head>. Also **all tags** are **not supported** by all World Wide Web browsers. If a browser does not support a tag, it will simply ignore it.

When you write HTML text, you can never be sure now the text is displayed in another browser. The text will **be reformatted every time** the user resizes his window. Never try to format the text in your editor by adding empty lines and spaces to the text.

HTML will **truncate the spaces** in your text. Any number of spaces counts as one. In HTML a new line counts as one space.

Using empty paragraphs <p> to insert blank lines is a bad habit. Use the <br> tag instead. You might have noticed that paragraphs can be written without the closing tag </p>. Don't rely on it. The next version of HTML might not allow you to skip any closing tags.

HTML automatically **adds an extra blank line** before and after some elements, like before and after a paragraph, and before and after a heading.

## 4.3 THE MINIMAL HTML DOCUMENT

Every HTML document should contain certain **standard HTML tags**. Each document consists of head and body text. The head contains the title, and the body contains the actual text that is made up of paragraphs, lists, and other elements. Browsers expect **specific information** because they are programmed according to HTML and SGML specifications.

### ➤ Simple html document

```
<HTML>
  < HEAD >
    <TITLE> HOME PAGE
    </TITLE>
  </HEAD>
  <BODY> THIS IS MY FIRST WEB PAGE
  </BODY>
</HTML>
```

(When you save an HTML file, you can use either the .htm or the .html extension. Save the file as "firstpage.htm". Open this file through your Internet browser.)

### Example Explained

The first tag in your HTML document is <HTML>. This tag tells your browser that this is the start of an HTML document. The last tag in your document is </HTML>. This tag tells your browser that this is the end of the HTML document. The text between the <HEAD> tag and the </HEAD> tag is header information. Header information is not displayed in the browser window. The title is displayed in your browser's caption.

The text between the <BODY> tags is the text that will be displayed in your browser.

The text between the <TITLE> tags is the title of your page

## 4.4 BASIC HTML TAGS

The most important tags in HTML are tags that define headings, paragraphs and line breaks.

### 1.<HTML>

The HTML tag identifies a document as an HTML document. All HTML documents should start with the <HTML> tag and end with the </HTML> tag.

Syntax

<HTML>.....</HTML>

e.g.

The following example begins and ends a short document with the HTML tag.

```
<HTML>
<BODY>
  This is HTML file.
</BODY>
</HTML>
```

Following tags appear in <HTML>tag:--

#### • HEAD

The HEAD contains general information, or *meta*-information, about the document. The HEAD tag defines an HTML document header. It is the first thing in any document, lying above the BODY and just after the <HTML> tag starting the document. The contents of the HEAD are not displayed as part of the document.

The HEAD tag can contain TITLE, BASE, ISINDEX, META, SCRIPT, STYLE, and LINK tags.

Syntax

<HEAD>...</HEAD>

e.g. <HTML>

```
<HEAD>
    <TITLE>WELCOME TO FIRST WEBSITE</TITLE>
</HEAD>
</HTML>
```

- **TITLE**

The title of a document is specified by the TITLE element, which should be placed in the document HEAD. Each document can have **only one title**, which should identify the document content in a general way.

The Title is **not part of the document text** and cannot contain hypertext links or special markup commands -- it must be simple text. Often the title is used to label the window displaying the text, or is used to label a place in a browser's history or bookmark list. It therefore should be short -- less than 64 characters  
Other tags can be placed are:

- **STYLE** -- Stylesheet instructions, written in a stylesheet language. Stylesheet instructions specify how the document should be formatted for display. Very few browsers currently support stylesheets.
- **SCRIPT** -- Script program code -- for enclosing, within a document, scripting program code that should be run with -- and that can interact with -- the document. Example languages are JavaScript and VBScript.

E.g. <HTML>

```
<HEAD>
    <SCRIPT language="VBSCRIPT"> VBscript </SCRIPT>
</HEAD>
<BODY>
    .... text of the document
</BODY>
</HTML>
```

## 2.<BODY> Defines the document's body

The BODY element contains **all the contents of a document**. Various mark-up elements are allowed within the body to indicate headings, paragraphs, lists, hypertext links, images, and so on.

### ➤ Attributes in body tag

#### 1) The BACKGROUND Attribute

This allows you to specify an image file to use as a background (a bit like a watermark) behind the displayed text and graphics.

E.g. <BODY BACKGROUND="c:\a.gif">

Text....

</BODY>

So image a.gif will be set as a background to your web page.

## 2) Background color of the web page

Attribute is: BGCOLOR="#rrggbb"

Sets the background color to the specified RGB color value, where RR GG and BB are the hexadecimal color codes for the Red, Green and Blue levels, ranging from 0 to 255 -- that is, 00 to FF. The color "000000" is black, while "FFFFFF" is white.

## 3) Setting the text color(TEXT Attribute)

Syntax: <BODY TEXT="#rrggbb">text in a body</BODY>  
Sets the default text color to the specified RGB color value.

### 4) Setting color for hyperlinks (LINK Attribute)

Syntax: <BODY LINK="#rrggbb">text in a body</BODY>  
Sets the default text color of hypertext anchors to the specified RGB color value.

## 5) Setting color for visited hyperlinks (VLINK Attribute)

Syntax: <BODY VLINK="#rrggbb">text in a body</BODY>

Sets the default text color of *visited* hypertext links to the specified RGB color value.

## ➤ Elements in the BODY are categorized as

### A) Text Block Elements

BODY element contains all the displayed content of a document. Structurally, the document content is organized into blocks of text, such as paragraphs, lists, headings, paragraphs, block quotations, and so on. These are generically called *block elements*, since they "block" chunks of text together into logical units. Block elements can often contain other blocks -- for example, a list item can contain paragraphs or block quotations, so that these elements can often nest together.

The block-level elements are:

- Hn (Headings) (h1 to h6)
- P
- ADDRESS
- BLOCKQUOTE
- PRE
- HR
- FORM
- TABLE

### B) Text Emphasis Elements

These are elements that mark text for special meanings, for example, that a particular piece of text is emphasized (EM) or a citation (CITE), or that specifies the desired physical formatting, such as boldface (B) or italics (I). These elements can usually appear anywhere inside a block element, with a few exceptions (you can't have images inside a PRE element).

### C) Special Elements -- Hypertext Anchors

Analogous to the text-level markup is the anchor (A) element. This is the element that marks hypertext links. Obviously you want to know a lot about this one.

### D) Character-Level Elements

Then are what I call character-level elements, namely line breaks (BR) and images (IMG). These are treated much like characters, and can appear wherever there is a character in a document.

### E) Character References

Finally there are character or entity references. These are special HTML "escape" codes that can be used to enter special characters that are hard to type, such as accented or other non-ASCII characters. You also need to use these to type angle brackets or ampersand characters -- as these are otherwise interpreted as HTML tags (< ... >) or as the beginnings of character or entity references (&).

Analogous to the text-level markup is the anchor (A) element. This is the element that marks hypertext links. Obviously you want to know a lot about this one.

### 3. <P> Defines a paragraph

Paragraphs are defined with the <P> tag. Can contain align attribute for alignment of the text within paragraph.

E.g. <P ALIGN="CENTER"> this is a paragraph</P>

HTML automatically adds an extra blank line before and after a paragraph. You must indicate paragraphs with <P> elements. A browser ignores any indentations or blank lines in the source text. Without <P> elements, the document becomes one large paragraph. The </P> closing tag may be omitted. This is because browsers understand that when they encounter a <P> tag, it means that the previous paragraph has ended. However, since HTML now allows certain attributes to be assigned to the <P> tag, it's generally a good idea to include it.

### • ADDRESS Element

The ADDRESS element is used for address information, signatures, statements of authorship, etc. It is often placed at the bottom (or top) of a document. The rendering of the contents of the ADDRESS is left up to the browser -- most browsers render the ADDRESS in italics. It may also be right justified, or indented.

e.g.  
 <ADDRESS><A HREF="c:\add.html">X.Y.Z.</A></ADDRESS> </P>  
 <P><ADDRESS>  
 WEB DESIGNER<BR>  
 Tel (023) 122 123.  
 </ADDRESS> </P>  
 These are rendered as

X.Y.Z  
 WEB DESIGNER  
 Tel (023) 122 123.

An address cannot contain P, BLOCKQUOTE, FORM or other block elements, but can contain text, text markup (emphasis, etc.), anchor elements or even images.

#### 4. <BR> Inserts a single line break

##### Line Breaks:

The <BR> tag is used when you want to end a line, but don't want to start a new paragraph. The <BR> tag forces a line break wherever you place it.

E.g. This <BR> is line breaks.

So here This and is line breaks will appear on two different lines.

The <BR> tag is an empty tag. It has no closing tag.

#### 5. <HR> Defines a horizontal rule.

Used to produce a horizontal line, the width of browser's window.

It allows to differentiate sections of your document.

SIZE and WIDTH attributes will let you alter the thickness

And the percentage of the windows covered by it.

e.g.<BODY>  
 <H1>This is my first web page</H1>  
 <HR SIZE=5 WIDTH="20%">  
 </BODY>

#### 6. <! Define a comment in the HTML source code Comments in HTML.

The browser will ignore a comment. You can use comments to explain your code, which can help you when you edit the source code at a later date.

E.g.<!-- This is a comment -->

Note that you need an exclamation point after the opening bracket, but not before the closing bracket.

Comments **do not nest**, and the double-dash sequence "--" may not appear inside a comment except as part of the closing --> tag. You must also make sure that there are no spaces in the start-of-comment string.

For example, the line

```
<!-- This is a valid comment -->
<!-- This is not a valid comment -->
```

is not, since there is a space between the left angle bracket and the exclamation mark.

## 7. Headings

**<H1>...<H6>** Defines heading 1 to heading 6.

Headings are defined with the **<H1>** to **<H6>** tags. **<H1>** defines the largest heading. **<H6>** defines the smallest heading (as shown).

Heading Size 1	Heading Size 2
Heading Size 3	Heading Size 4
Heading Size 5	Heading Size 6

E.g.

```
<H1 STYLE="COLOR: blue">THIS is displayed in large font with blue color </H1>
<H1 STYLE="FONT-FAMILY:verdana">This is displayed in verdana font style
</H1>
<H1 STYLE="FONT-SIZE:150%">You can define size in percentage also.</H1>
HTML automatically adds an extra blank line before and after a heading.
```

## 8. <PRE> tag

This tag preformats the text. The text appearing between **<pre>** and **</pre>** is displayed in monospace form. Using this tag, we can position the characters. It can also be used for columnar lists.

Eg. **<PRE>** Employee\_name Employee\_number Employee\_address **</PRE>**  
So output is displayed as :

Employee\_name      Employee\_number      Employee\_address

## 4.g. FONT TAGS IN HTML

To have different size and color to the text, rather than using headings tags, many people uses **<FONT>** tag.

The **<FONT>** tag in HTML is deprecated. It is supposed to be removed in a future version of HTML.

### Font attributes:

Attribute	Example	Purpose
size="number"	size="2"	Defines the font size
size="+number"	size="+1"	Increases the font size
size="-number"	size="-1"	Decreases the font size
face="face-name"	face="Times new Roman"	Defines the font-name
color="color-value"	color="#eef00	Defines the font color
color="color-name"	color="red"	Defines the font color

e.g.  
<P>  
<FONT SIZE="5" FACE="VERDANA" COLOR="#AA00FF">  
This is example of demonstrating fonts. </FONT>

</P>

- **<MARQUEE> TAG**  
MARQUEE, supported only by the Microsoft Internet Explorer 2 (and later) browser, is used to create a scrolling text marquee.

e.g. <MARQUEE ALIGN="top">Scrolling text </MARQUEE>  
creates a text marquee with the enclosed text scrolling along the frame.

The another attribute is DIRECTION defining direction of the marquee text.

e.g. <MARQUEE DIRECTION="RIGHT">WELCOME</MARQUEE>

So here WELCOME scrolls from left towards right. The default direction is right to left.

But other commercial browsers, including Netscape Navigator, do not support this element.

### ➤ Character Formatting

HTML has two types of styles for individual words or sentences: logical and physical. Logical styles tag text according to its meaning, while *physical styles* indicate the specific appearance of a section. For example, in the preceding sentence, the words "logical styles" was tagged as "strong." The same effect (formatting those words in italics) could have been achieved via a different tag that tells your browser to "put these words in bolds." Try to be consistent about which type of style you use.

#### Logical Styles:

<EM>  
for emphasis. Typically displayed in italics. (*Consultants cannot reset your password unless you call the help line.*)

<CODE>  
for computer code. Displayed in a fixed-width font. (The <stdio.h> header file)

<STRONG>  
for strong emphasis. Typically displayed in bold. (**NOTE:** Always check your links.)

<VAR>  
for a variable, where you will replace the variable with specific information.  
Typically displayed in italics. (rm *filename* deletes the file.)