2.6 Formatting the HTML Text

Text plays predominant role in Web pages. Reader's eyes should not turn to red when he looks at the text in the Web pages. For that the text must be well aligned and formatted to ensure readability. HTML provides many elements to attain the same. Let us now discuss them in detail.

The Meaning Elements

The "meaning" elements are used to identify words or phrases that have meaning to the text. These elements show characteristics above the text. For example, a person's name would be identified by the <PERSON> element, and the title of a book might be identified by the <CITE> element. These elements enhance the quality of your documents by adding structure to the document. Indexing tools and programs that you write can use these elements to identify pieces of the text that you as the author found important. When publishing documents on the Web, where data is valuable, representing a document in a structured manner makes all the difference in the world

The <EM> (EMPHASIS) Element:

provides emphasis and is typically displayed as italics. The <EM> element should be used where the word or phrase should be emphasized rather than italicized.

The <CITE> Element:

is to denote a citation. The citation is usually displayed in italics.

The <STRONG> Element:

denotes strong emphasis. <STRONG> is usually displayed as bold

The <CODE> Element:

denotes that the text is some type of computer output or example of programming code. <CODE> is usually displayed in a fixed-pitch font. The same font is usually used with <PRE>; however <PRE> is for preformatted text while the browser will format the text associated with <CODE>

The <SAMP> Element:

denotes characters that are to used as sample text

The <KBD> Element:

denotes text that might be typed at the keyboard by a user. Of course, this is useful when writing books like this one or for writing instructions to be followed

The <VAR> Element:

denotes the text is a variable name as one might find in a C or Perl program

The <DFN> Element:

denotes the definition of a term

The <Q> Element:

is used to denote a short quotation

The <AU> Element:

is used to identify an author

The <PERSON> Element:

is used to identify the names of people

The <ACRONYM> Element:

denotes the word or phrase to be an acronym

The <ABBREV> Element:

is used to denote abbreviations

The <INS> Element:

denotes text that has been inserted into the surrounding text. This is handy when trying to show differences to a document, much like change bars are used in some work processors

The <DEL> Element:

denotes text that has been deleted from the surrounding text. This is handy when trying to show differences to a document, much like change bars are used in some word processors

Example:

```
<html>
<head>
        <title>Logical Formatting Examples</title>
</head>
<body>

The <EM> spirit of the law </EM> is sometimes more <EM> important </EM> <p>

To be or not to be", <CITE>Shakespeare</CITE> <p>
```

When coming to a red light at an intersection, <STRONG> you must stop </STRONG> <p>

One of the first lines that every C programmer learns is: <CODE> puts (“Hello
 World!”); </CODE> <p>

<SAMP><H1>This is the first headline</Hl></SAMP> <p>

At the prompt enter your login name. <p>
Login: <KBD>farrar </KBD> <p>
At the next prompt enter your password. Your password will not be displayed on the screen <p>
Password: <KBD>!!urdead</KBD> <p>

In the above perl example, the <VAR> $height </VAR> is a constant, while <VAR> $width </VAR> increments by 5<p>

<DFN> Aqua </DFN> is typically known as water<p>

<Q>To be or not to be? That is the question </Q> <p>

</body>
</html>

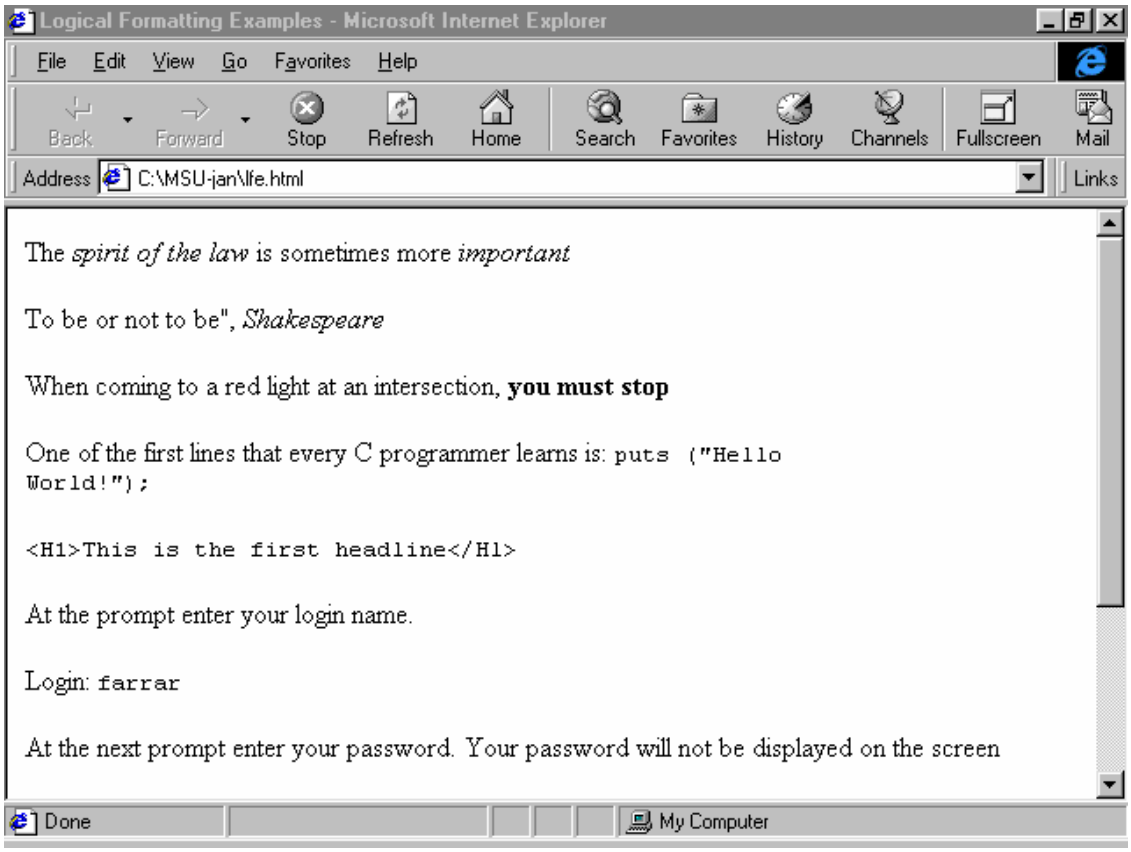

Figure 2.7 *Example for The Meaning Elements*

The Look Elements

The "look" elements deal with how the text is to be displayed not its meaning. These are useful when you want to italicize or bold a word or phrase

The <B> (BOLD) Element:

is used to denote boldface. Bolding characters make them stand out in the text.

The **<I>** (ITALIC) Element:
is used to denote italics

The **<TT>** (TELETYPE) Element:
is for teletype or typewriter style fixed-pitch font

The **<U>** (UNDERLINE) Element:
is used for underline. Underline is widely supported by most browsers, but some ignore the <U> element or display the text as italics

The **<STRIKE>** Element:
is used for showing strikethrough. This is handy when you want to show a phrase is no longer needed, while still allowing the text to be seen

The **<BIG>** Element:
is used to display the character, word, or phrase in a larger font compared to the rest of the text

The **<SMALL>** Element:
is used to display the character, word or phrase in a smaller font compared to the rest of the text

The **<SUB>** Element:
is for showing subscripts (those below the line)

The **<SUP>** Element:
is for showing superscripts (those above the line)

Example

```
<html>
<head>
        <title>Look Elements</title>
</head>
<body>
This is in <B> Bold </B> text<p>
This is in <I> Italics </I> text<p>
Example for <TT> teletype </TT> text<p>
Example for <U> Underlined </U> text<P>
```

This text has been <STRIKE> striked</STRIKE><p>
<BIG>O</BIG>nce upon a time, in a land not so far away, in a time not so long ago..<p>
<SMALL>O</SMALL>nce upon a time, in a land not so far away, in a time not so long ago…<p>
H₂O is the chemical designation of water<p>
Pi * r² is the mathematical formula for the area of a circle
</body>
</html>

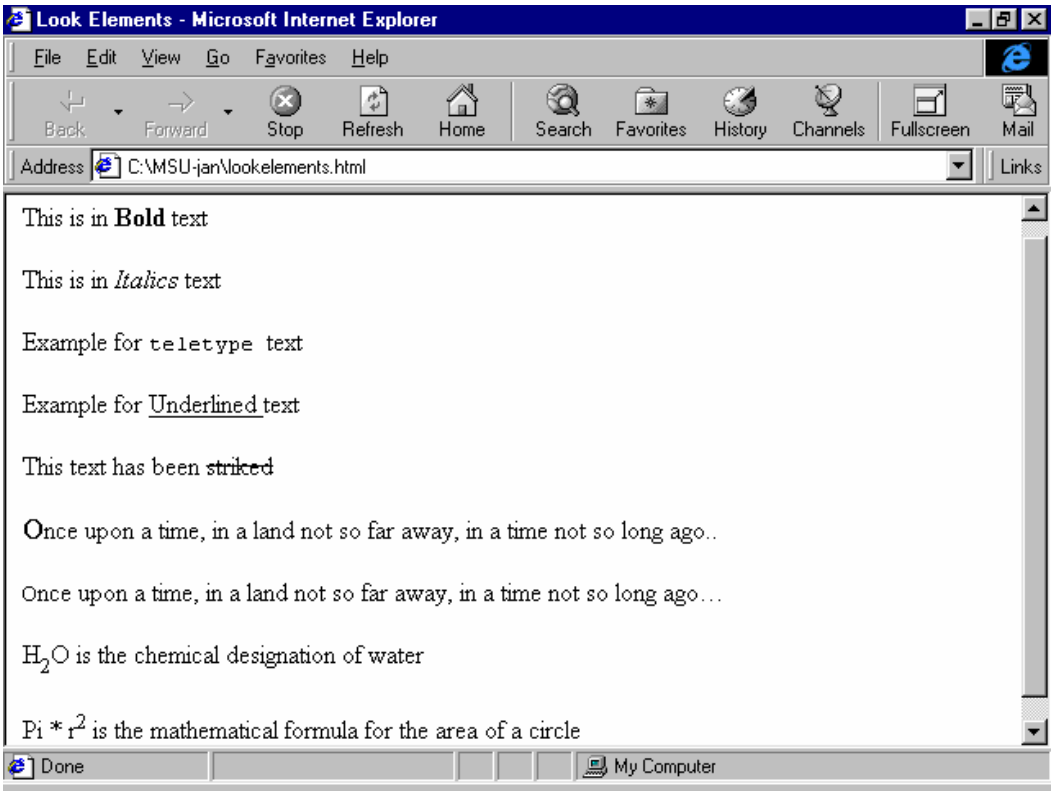

Figure 2.8 *Example for the Look Elements*

Preformatted Text, Spaces and All: <PRE>

The method for making several lines of text appear in the way they are typed is to mark the whole section as preformatted text using the <PRE> element.

Syntax

<PRE> …*Text*…</PRE>

Text is any text, including returns, spaces, and other hard formatting. See the following listing and Figure 2.9. All the text is contained by the <PRE> tags, which causes the browser to display everything in between exactly as it finds it. It also uses a mono-spaced font for display, which helps format the text into rows and columns for data presentation.

```
<HTML>
<HEAD>
<TITLE>Hot Water Safety</TITLE>
</HEAD>
<BODY BACKGROUND="D:\Raguram\Images\backgrd.gif">
<B>
<PRE>
Hot water is dangerous.

  Turn down the hot water heater thermostat.

    Always supervise children in the bathtub.

      Don't encourage children to play around the tub.
</PRE>
</B>
</BODY>
</HTML>
```

Figure 2.9 *Appearance of the preformatted text.*

Notice in the above Figure 2.9 that the mono-space font preserves character spacing for accurate display of indents and other formatting.

The mono-spaced font helps when preserving the line breaks and spaces inserted by the user. If he start a new line, insert five spaces, and then start typing, that's exactly how the text appears when using the <PRE> tag. Browsers disable automatic word wrap to further ensure that the text is displayed exactly as it is typed.

For browsers that aren't compatible with tables (explained in the next Chapter), preformatted text is an effective way to display a table made out of textual elements representing the border lines, such as the following:

```
<PRE>
============================
Country        |  Gold Medals
============================
India          |     199
---------------+--------------
Russia         |     101
---------------+--------------
USA            |      99
---------------+--------------
Japan          |      30
---------------+--------------
China          |      01
============================
</PRE>
```

Address Information: <ADDRESS>

The <ADDRESS> tag is used to mark contact information for the current document, whether it's an email address or a complete mailing address and phone number. It behaves much like a paragraph tag, forcing the text within its confines to be separated from surrounding material *by additional line spacing*.

Syntax

<ADDRESS>…*ContactInformation*…</ADDRESS>

ContactInformation is the address information, along with any paragraph-level formatting such as line breaks. It's typically displayed as italic body text as in Figure 2.10.

```
<ADDRESS>
Godzilla <BR>
 -9182/~ Rocket Highway<BR>
Moon<BR>
(605) 374-5716
```

Drawing a Line on the Page: <HR>

Horizontal lines are an easy-to-use element for dividing a page into logical sections. They signal the reader to be alert for a change in subject or style or they can separate figures and captions from body text. This tag also includes several attributes to fine-tune its appearance.

</ADDRESS>

Figure 2.10 ***Text formatted with <ADDRESS>.***

<HR [WIDTH=*lineWidth*] [SIZE=*lineThickness*] [NOSHADE]>

Here, *lineWidth* and *lineTickness* set the basic appearance of the line. The height of a horizontal rule is set in pixels, such as SIZE=6. The WIDTH is set either in pixels or as a percentage of the browser window width. If percentage value is used, it must be suffixed with a percentage sign (%) after the value. The typical defaults for height and width are 3 pixels high and 100 percent of page width (Figure 2.11.)

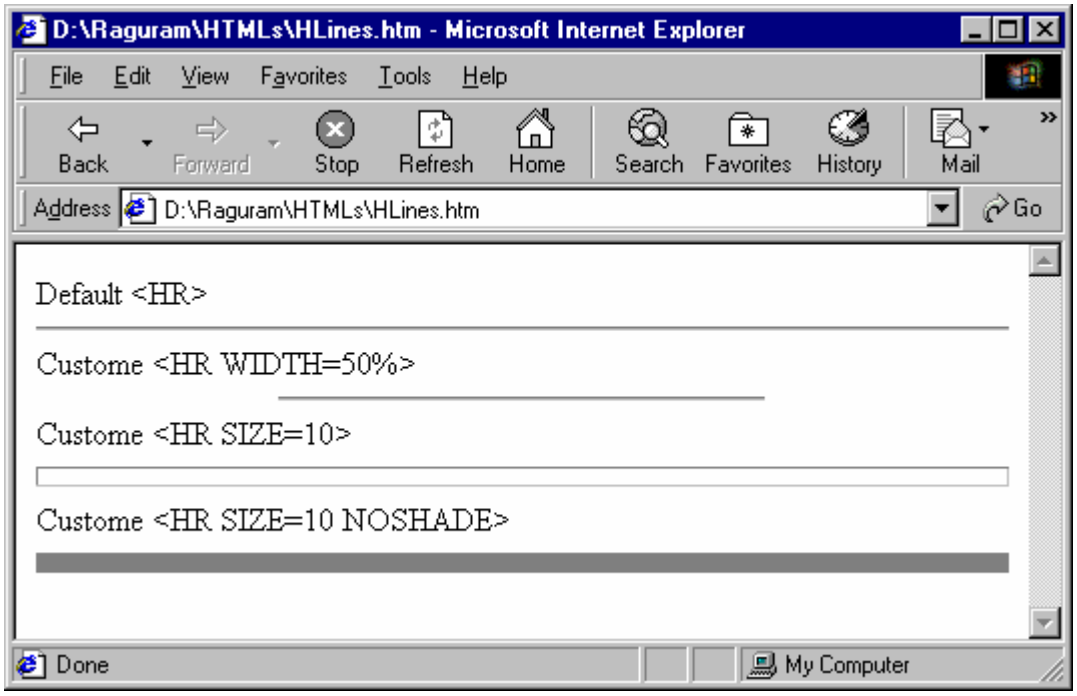

Figure 2.11 *Different styles of Horizontal Lines.*

The last attribute is NOSHADE. Normally, a browser displays a horizontal rule in some form of three-dimensional shading. This shading varies from browser to browser, some show it as depressed, some raised, some as on outline. Using NOSHADE forces the horizontal rule appear as a thick line with no additional highlighting. Again, this appearance can very among browsers. Some use black for the rule and others use shades of gray.

Centering a Block: <CENTER>

Text or embedded objects like images that are enclosed within <CENTER> and </CENTER> are centered in the browser according to its current coordinates. The <CENTER> element is simply an alias for <DIV ALIGN= "CENTER"> and is treated exactly the same way. It is unlikely that the <CENTER> element will go away, considering its simplicity and widespread use.

Syntax

<CENTER>…Text or images </CENTER>

Dividing document into Sections: <DIV>

The <DIV> element is used to structure HTML documents into unique sections or divisions.

Syntax

```
<DIV [ALIGN = "LEFT" | "RIGHT" | "CENTER"]
      [ STYLE =  style information ]>

         … Text or images …

</DIV>
```

The ALIGN attribute makes it possible to align a portion of the document to the left, right, or center. By default, content within the <DIV> element is left aligned. Divisions are much useful when they are used with the STYLE attribute. Because this attribute enables to specify segments position, appearance as in below:

STYLE = "Position: Absolute; Left:20; Top:30; Font-weight: Bold; Height:100; Width:100; background-color:Yellow;color:Black"

The first property-value pair, Position:Absolute, tells the browser to position the DIV (segment) with respect to the top edge and left edge of the browser window.

Another possible value of for Position is Relative. In contrast with Absolute, Relative situates the segment in relation to other segments on the page.

The properties left and top specifies the x and y coordinates of the segment. The Height and Width properties species the size of the segment.

The following listing, which uses three <DIV> elements, produces the document as in Figure 2.12

```
<HTML>
<HEAD>
<TITLE>Divisions</TITLE>
</HEAD>

<BODY BACKGROUND="D:\Raguram\Images\backgrd.gif">

<CENTER>
Little Things                    <BR>
<I> that make a </I>             <BR>
<BIG><BIG>BIG</BIG></BIG>        <BR>
<I> Difference </I>
</CENTER>

<DIV ALIGN="Left" STYLE= "Position:Absolute;Left:20;Top:120;Height:80; Width:120;
background-color:Yellow;color:Black" >
You can win more friends by your ears than by your tongue.
</DIV>

<DIV ALIGN="Center" STYLE= "Position:Absolute;Left:190;Top:120;Height:80; Width:120;
background-color:Yellow;color:Black" >
Your children are mirrors - the reflection of what you live.
</DIV>


<DIV ALIGN="Right" STYLE= "Position:Absolute;Left:360;Top:120;Height:80; Width:120;
background-color:Yellow;color:Black" >
The best way to destroy your enemies is to make them friends.
</DIV>

</BODY>
</HTML>
```

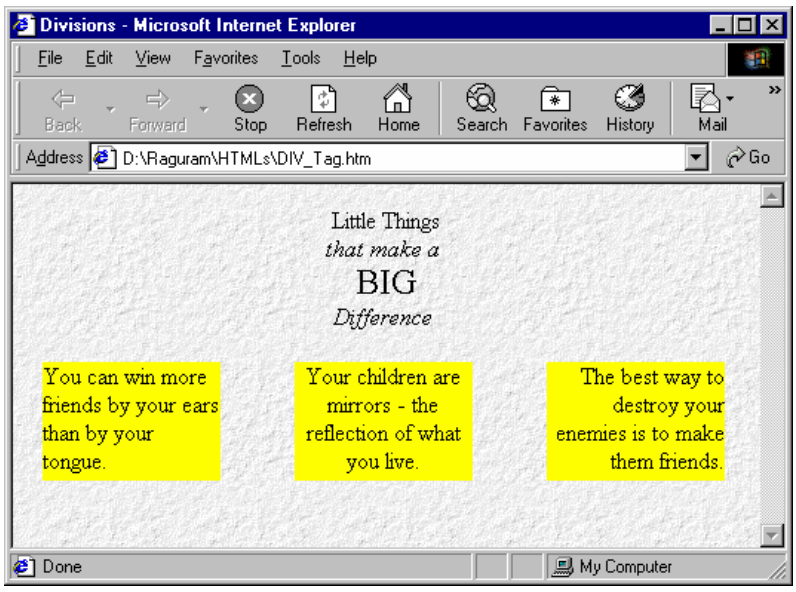

Figure 2.12 *Appearance of the document that uses Divisions.*

Character Entities

After covering the basic text formatting elements, there would appear to be nothing left to talk about – but there is still one more level to HTML documents: the characters themselves.

Sometimes it is necessary to put special characters within a document. These include accented letters, copyright symbols, or even the angle brackets used to enclose HTML elements. To use such characters in an HTML document, they must be "escaped" using a special code. All character codes take the form *&code;* where code is a word or numeric code indicating the actual character that is to be put on the screen. Some of the more commonly used characters are shown in the following table.

Numeric Value	Named Value	Symbol	Description
&#034;	&quot;	"	Quotation mark
&#038;	&amp;	&	Ampersand
&#060;	&lt;	<	Less than
&#062;	&gt;	>	Greater than
&#153;	N/A	™	Trademark
			Non-breaking space
&#169;	&copy;	©	Copyright symbol
&#174;	&reg;	®	Registered trademark

A Few common character entities.

The following listings which uses character entities produces the document as in Figure 2.13

<HTML>
<HEAD>

```
<TITLE>Character Entities</TITLE>
</HEAD>

<BODY>

<H1 ALIGN="Center">Character Entities</H1> <HR>
<P>
Character entities like &amp;copy; allow users
to insert special characters like &copy;. <BR><BR>

One entity is often abused: the non breaking space <BR>

Inserting spaces is easy with &amp;nbsp;<BR>

Look:   S       P      
A       C      
E       S. <BR>

</P>
<HR>
<ADDRESS>
Contents of this page
&copy; 2000 Character Corporation  <BR>
The <B>Character</B> &#153;
a rigistered trademark of Star Company, Inc.
</ADDRESS>
 </BODY>
</HTML>
```

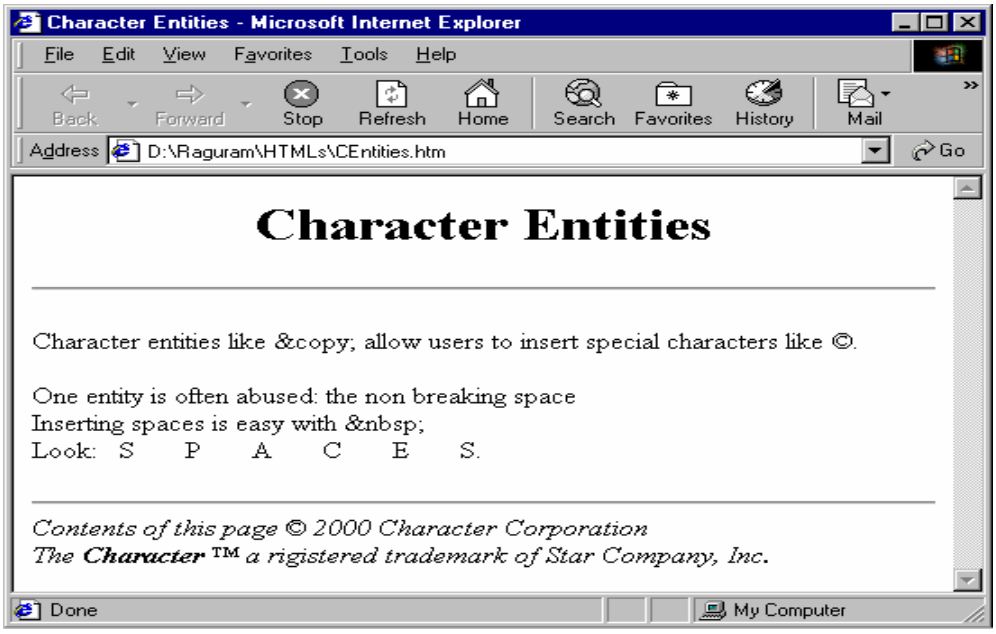

Figure 2.13 *Rendering of character entities.*

Commenting the Lines: <!--…-->

While designing large web pages, the designers may get confused even with his own sentences. This can be avoided by adding meaningful comments. HTML provides tags to make comments whose syntax is given below:

<!-- … -->

Example

<!-- This is an informational comment that can occur

anywhere in an HTML document -->

HTML Rules (For Self Reading):

There are some rules to remember when writing HTML. These are briefly reviewed here.

- **Element names are not case sensitive.** An element like <hTml> is equivalent to <html> or <HTML>. Element case does not matter to a browser. However, writing elements consistently in upper – or lowercase makes HTML documents easier to understand and maintain. Convention suggests that uppercase is the preferred practice.
- **Attribute names are not case sensitive.** Just as <body> is equivalent to <BODY>, <BODY BGCOLOR= "Yellow"> is equivalent to <BODY bgcolor= "Yellow". As with elements, consistent use of case improves legibility, and uppercase is preferred.
- **Attribute values may be case sensitive.** The value of an attribute may be case sensitive, especially if it refers to a file. The file name in <IMG SRC= "filename.gif"> may not be the same as the filename in <IMG SRC= "FILENAME.GIF">; it depends on whether or not case matters to the operating system of the server containing the file. For best results, always specify a filename exactly as it has been saved.
- **Element names cannot contain spaces.** Browsers treat the first space encountered inside an element as the end of an element's name and the beginning of its attributes. For example, <I M G> does not mean <IMG>, the image element. It means <I>, the italic element, with two undefined attributes M and G.
- **Browsers collapse and ignore space characters in HTML content.** Browsers collapse any sequence of spaces, tabs, and returns in an HTML document into a single space character. These characters convey no formatting information, unless they occur inside a special formatting element like <PRE> which preserves their

meaning. Extra spacing can be used liberally within an HTML document to make it more legible to HTML authors.

- **An element that encloses the start tag of another element must also enclose its end tag, if one exists.** Elements often contain other elements inside the document section they enclose. Any element that starts within a section enclosed by another must also end there. In other words, an element's tag pairs should be nested within one another and their end tags should no cross. To make some text bold and italic, use <B><I>Correct</I></B> and not <B><I>Not Correct</B></I>.
- This is primarily a stylistic issue, since no major browsers at this time have problem with this. Authors are still advised to nest tags rather than cross them. Incorrect nesting may result in incompatibilities with emerging technologies, and may cause rendering problems in some esoteric browsers.
- **Browsers ignore unknown elements.** Browsers ignore elements they do not understand. They do attempt to interpret any content enclosed by an unknown element. If a browser does not understand the <STORY> element in <STORY> Titanic – A Tale of Love </STORY>, it ignores it. It does, however, render the words "Titanic – A Tale of Love" as normal text.
- **Browsers ignore unknown attributes.** As with elements, browsers ignore any attributes they do not understand.

2.7 Short Summary

- The web browsers in use – NCSA Mosaic, Netscape and Internet Explorer
- A hypertext document is an electronic document that contains links to related pieces of information.
- Hypertext is a nonlinear way to access information
- HTML (Hypertext Markup Language) is the language that puts the face on the Web by helping to prepare documents for online publications.
- The collection of HTML pages makes up the *World Wide Web*
- HMTL is Not a WYSIWYG Design Language
- An HTML file is enclosed within <HTML> and </HTML> elements, which indicates that the contents of the file include markup.
- The "*meaning*" elements are used to identify words or phrases that have meaning to the text.
- To put special characters within a document, we use the Character Entity

2.8 Brain Storm

1. What is a Web browser? Discuss the browsers available.
2. Hypertext is a nonlinear way to access information – Discuss.
3. Is HTML a programming language?
4. List down the benefits of HTML
5. What is an HTML Tag (Element)?
6. Discuss the parts of an HTML document
7. What is the significance of using a Head Tag in an HTML Document?
8. What is the Tag used for setting the background color for an HTML Document?

   a. BG color b. Background color c. Bgrnd color d. Backcolor
9. List down the various attributes of body tag.
10. What are the ways of giving values for color?

    a. colorname
    b. RRGGB format
    c. RGB combination
    d. None of the above
11. What is the range for the Headings Tag?
12. What is the difference between the output obtained using the <P> and
 Tags?
13. Differentiate between the Meaning Elements and Look Elements.
14. What is the use of <ADDRESS> Tag? How it differs from <I> Tag?
15. What is the significance of using <DIV> tag?
16. HTML Elements are case sensitive – True or False?

Lab Unit (2 Real Time Hours)

At the end of this exercise you would have completed building your WEB page.

-Open the Notepad and give the necessary TAGS required for a WEB Page.

-The WEB Page is to display Information about RADIANT in the following way:
-The page should have a title 'Radiant Home Page'.

-The heading should be 'RADIANT SOFTWARE LIMITED'.

-The contents of the page:

RADIANT the pioneer in higher end technology training and specializes in Client Server Computing, GUI and OOPS methodology and Dedicates itself to evolve a New Era in Software Education.

Radiant Provides Job Oriented training with Ultimate Technical Excellence.

CITE Offers the Following Courses:

Oracle

Visual Basic
Oracle DBA
Unix Administration

CADRe

Unix C,C++

The courses are offered in the following timings:

Part Time
Full Time
Crash

Note:

- **Try using various Heading styles.**
- **Give a Back color to the page and try changing the back color.**
- **Try using various font sizes, colors and font faces.**
- **Use your own data and text formatting to make this page more attractive, informative and impressive.**