



Introduction to HTML

Working of the Internet and HTML

Computers are widely used in variety of applications. When computers are connected with each other they can share resources. Such group of connected computers is known as a computer network. The Internet is collection of such multiple computer networks, hence known as network of networks. On the platform of the Internet, a distributed information system exists, which is called World Wide Web, WWW or Web in short. The notion of the Web was conceived in 1991 by Tim Berners-Lee (figure 1.1), while consulting at CERN (a European Organization for Nuclear Research, <http://cern.web.cern.ch/CERN/>) in Switzerland. The Web is a repository of multimedia information on the internet platform. The web content in form of web pages is explored using browsers (special applications to retrieve and view web information). On these web pages, links are placed pointing towards different locations. These links are known as hyperlinks. Clicking on such hyperlink, one can redirect himself to an intended location. This operation is known as following the hyperlink. Any content such as text, picture, graphics, etc. can be embedded with such hyperlink.

The content and hyperlinks cannot be directly expressed on the Web. Hyper Text Markup Language (HTML), is needed to describe how a web page should be displayed by a web browser. Thus the HTML is considered as a language for describing web pages. The HTML is a documentation language to mark content of web pages such as heading, title, table, image, etc. It is machine independent and all Internet browsers accept the content written using HTML code.



Figure 1.1 : Tim Berners-Lee

HTML is a kind of markup language. A markup language is a set of tags that enables additional information (besides the content) on how to present the web content. HTML files are text files that contain additional formatting markup information in form of tags along with its content. The HTML is the most popular markup language; and it offers fixed set of tags. HTML is derived from of SGML (Standardized General Markup Language), which was developed by the International Organization for Standards (ISO) in 1986 to facilitate the sharing of machine-readable documents.

An HTML code is thus a combination of content to be displayed on a web page using browser and tags that helps in guiding the presentation of the content. Without such building block codes, it is impossible to display content on web pages. This makes HTML coding compulsory utility for web page creation, interpretation and presentation.

A Simple HTML Document

Let us create a simple web page that discusses about rainbow using HTML. The contents of the web page are shown in table 1.1.

RAINBOW

Rainbow consists of seven colours. These colours are Violet, Indigo, Blue, Green, Yellow, Orange and Red. They are also acronymed as VIBGYOR.

Rainbow is caused by reflection of light in water droplets in the Earth's atmosphere, resulting in a spectrum of light appearing in the sky. It takes the form of a multi coloured arc.

Table 1.1 : Text to be displayed on web page using HTML

HTML code to display the contents shown in table 1.1 is given in code listing 1.1.

```
<html>
  <head>
    <title> About Rainbow
  </title>
</head>
<body>
  <h1> RAINBOW </h1>
  <p>  Rainbow consists of seven colours. These colours are
    Violet, Indigo, Blue, Green, Yellow, Orange and Red.
    They are also acronymed as VIBGYOR.
  </p>
  <p>  Rainbow is caused by reflection of light in water
    droplets in the Earth's atmosphere, resulting in a
    spectrum of light appearing in the sky. It takes the
    form of a multi coloured arc.
  </p>
</body>
</html>
```

Code Listing 1.1 : Sample HTML Code

Observe that besides the content about rainbow, code listing 1.1 also displays several sets of angular brackets with words or letters within them. These brackets and words inside them are known as tags.

A tag is made up of letters, words and numbers enclosed between a left and right angular bracket.

A combination of opening and closing tag along with some content between the two tags forms an element. An HTML element may be empty or can have some attributes to specify the additional formatting and publishing instructions. Figure 1.2 illustrates structure of tags and elements with an example.

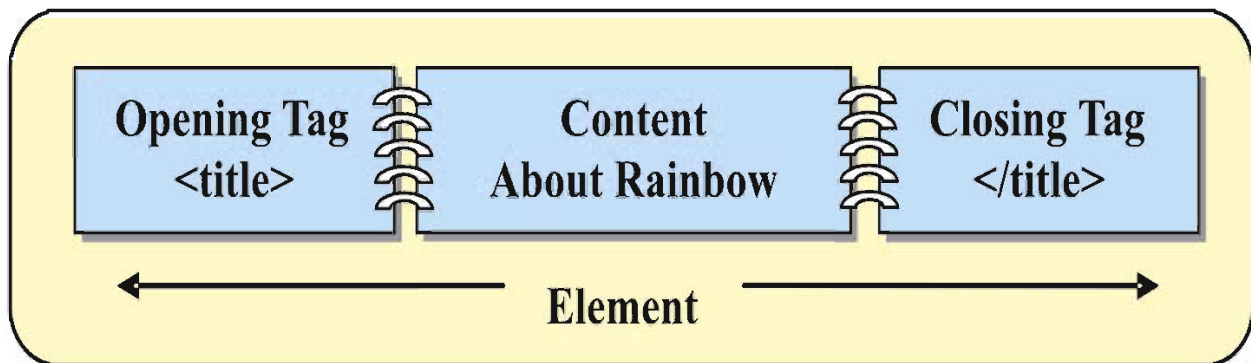


Figure 1.2 : An HTML element

Figure 1.2 indicates that `<title> About Rainbow </title>` forms an element of the HTML code. Observe that a closing tag contents are preceded with a forward slash (`</title>`). It marks the end of an element. Together, the pair of tags and the content within them forms an HTML element. The title tag here defines title of the web page generated by the HTML code.

Another example of such element is as follows :

`<h1>` **RAINBOW** `</h1>`

Here h1 tag refers to heading.

As stated above, content available between `<h1>` and `</h1>` is identified as heading and presented as heading. Similarly, content available between `<p>` and `</p>` is identified as paragraph and presented as paragraph text. The whole document is embedded between opening `<html>` and closing `</html>` tags.

To view how this page will look in a browser, follow the given steps :

Step 1 : Open gedit editor using Applications → Accessories → gedit. The gedit is a general purpose text editor for the GNOME (part of a project called GNU, free software by MIT) desktop environment, Mac OS X and Microsoft Windows. Alternatively you may use a shortcut available for the gedit editor at the header row of the screen.

Step 2 : Type the HTML contents of code listing given in table 1.2 in the empty gedit Window. Figure 1.3 show the look of the gedit editor after you have typed the code. Save the code as "p1.html", by selecting save option shown at the header row of the editor. Note that the HTML file can be saved with html or htm extensions. Figure 1.4 shows method to save the code.

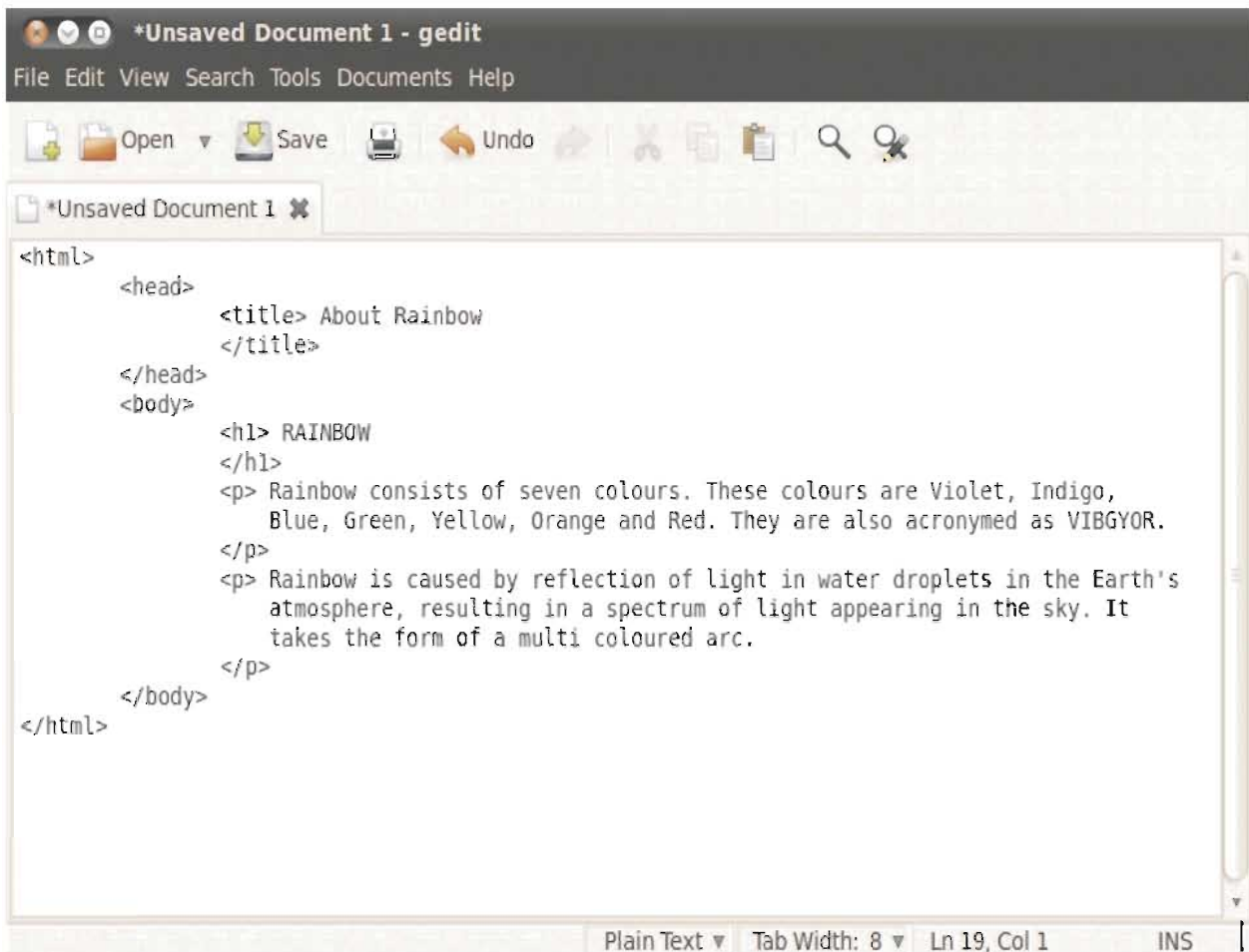


Figure 1.3 : The HTML code written in gedit editor

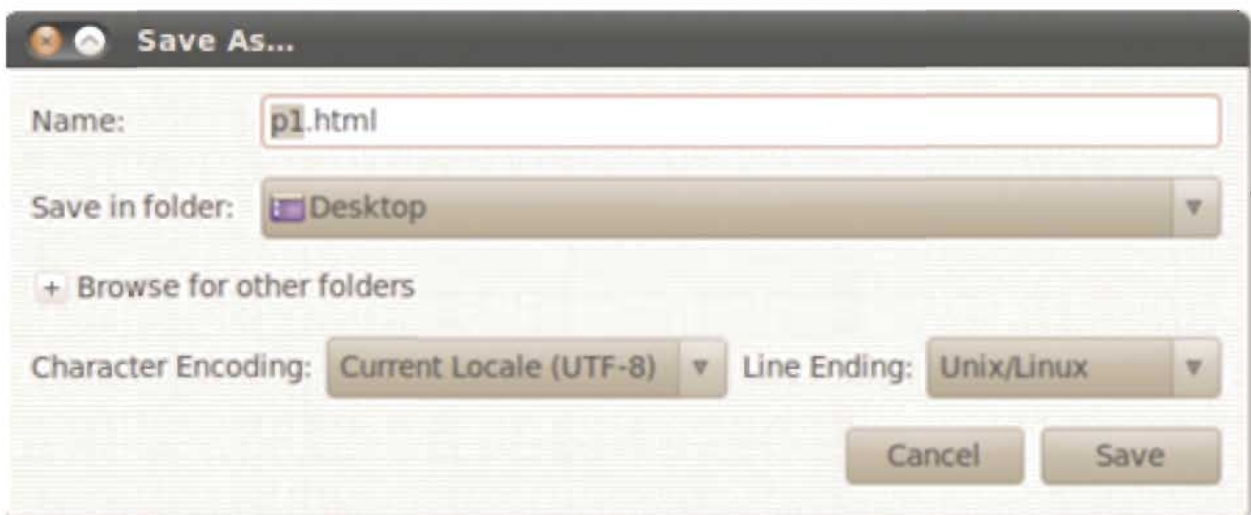
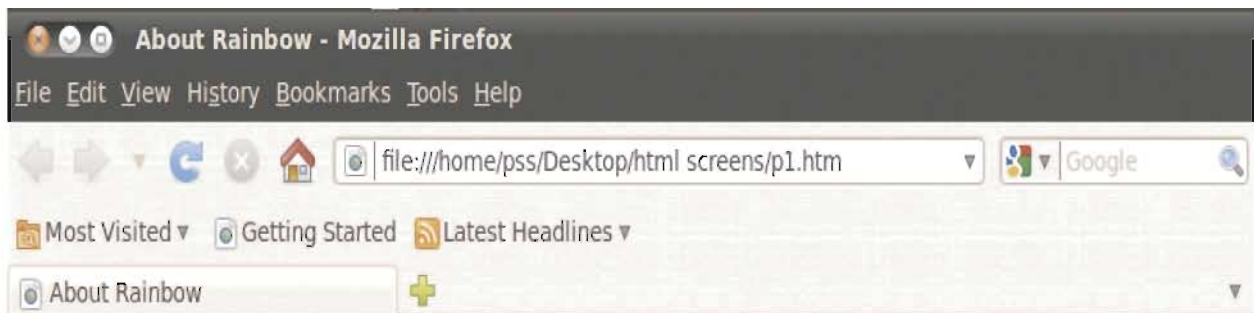


Figure 1.4 : Save As... dialog box

Step 3 : Open a browser such as Mozilla Firefox or any other browser that is installed on your computer. Select File → Open File, you will see an open file dialog box. Select the file that you want to open and click on Open button. Alternatively, you may double click on the file. Figure 1.5 shows the look and feel of the file when opened in Mozilla Firefox.



RAINBOW

Rainbow consists of seven colours. These colours are Violet, Indigo, Blue, Green, Yellow, Orange and Red. They are also acronymed as VIBGYOR.

Rainbow is caused by reflection of light in water droplets in the Earth's atmosphere, resulting in a spectrum of light appearing in the sky. It takes the form of a multi coloured arc.

Done

Figure 1.5 : The web page generated through the HTML code in a browser

It is to be noted that the tags used in HTML code are not case sensitive. The tags may be written in uppercase letters, lower case letters or mixture of upper and lower case letters. Further, multiple elements can be written in a single line. However, it is advisable to write each element in a new line with proper indentations for the sake of readability.

Structure of an HTML Document

HTML document is structured into two major parts. The first part is head section and second part is body section. They are also known as head element and body element. The head section contains information about the page such as title and description of the page. All these information should be embedded within the `<head>` and `</head>` tags.

The body element is embedded within the `<body>` and `</body>`. This is the content which can be seen within the browser. Both the head and body elements are embedded within the `<html>` and `</html>` tags.

HTML Title

The title of a web page is specified by the TITLE element, which should be placed in the head section of the document. It is to be noted that a document should have only one title element. It is used to identify the document content in a general way. Further, the content of title is not a part of the document text. Because of this, it should be simple text and cannot contain special commands such as hyperlinks. The title appears as a label of the window displaying the text. The title also holds a place in a browser's history or bookmark list. It is therefore recommended that title should be short. In the example HTML code given in listing of table 1.2, the title is "About Rainbow". It appears at the top of the windows displayed as shown in figure 1.5.

HTML Heading Style

HTML document generally begins with heading. In the example shown in code listing 1.1 heading style 1 (h1) tag is used. Observe the heading RAINBOW shown in the figure 1.5 to visualize how the heading style 1 looks. There are five more heading styles available in HTML. Heading can be

created in total six inbuilt sizes named as h1, h2, h3, h4, h5 and h6. These six levels of headings are described in the HTML code given in figure 1.6.



```
<html>
  <head>
    <title> About Rainbow
  </title>
  </head>
  <body>
    <h1> Rainbow </h1>
    <h2> Rainbow </h2>
    <h3> Rainbow </h3>
    <h4> Rainbow </h4>
    <h5> Rainbow </h5>
    <h6> Rainbow </h6>
  </body>
</html>
```

Figure 1.6 : HTML code for different levels of heading in HTML

The code is entered using the gedit editor. Save the code as "p2.html". When we see the code in a browser, it presents a view as shown in figure 1.7.

Most browsers display the contents of the <h1>, <h2>, and <h3> elements larger than the default size of text in the document. The content of the <h4> element is similar to the default content size. However, one can always redefine the sizes of these headings.

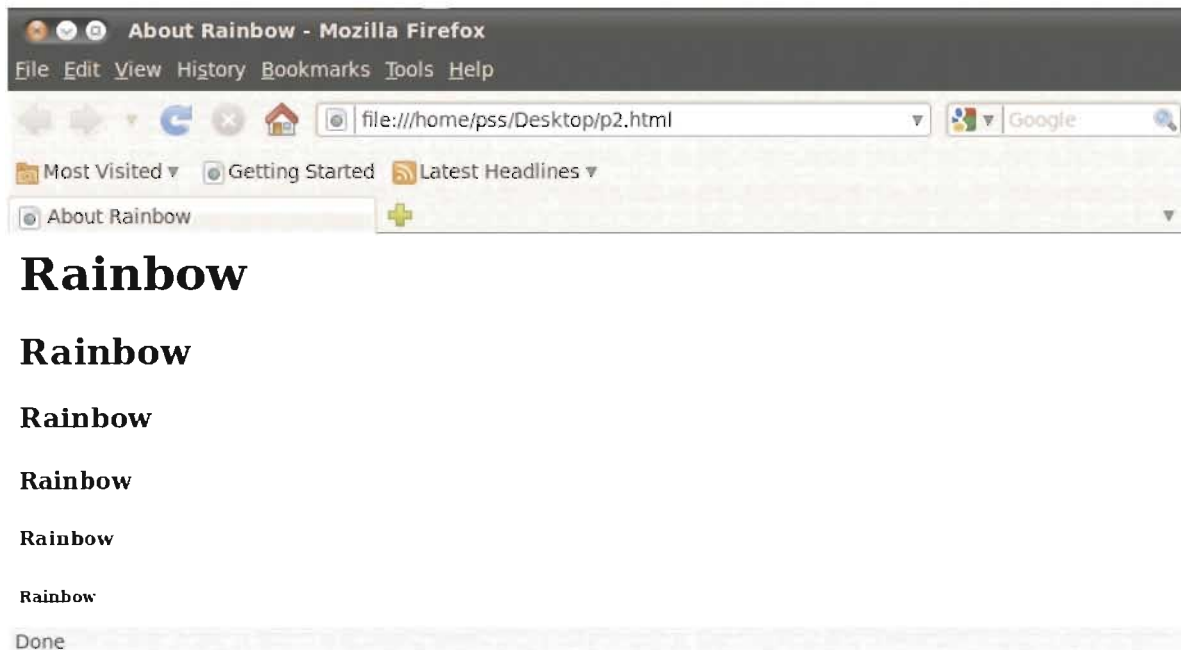


Figure 1.7 : Output of different levels of heading in HTML

Basic Text Formatting Tags

Besides the heading styles such as h1 to h6, there are more basic formatting tags available in HTML, which can be used in body section of an HTML code. Some of the useful tags are as follows :

Paragraph tags : <p> and </p>

The <p> tag structures the content into a paragraph. Each paragraph of text should go in between an opening <p> and closing </p> tag. Following are three valid examples of the same :

<p> This is first paragraph. </p>

**<p> The second paragraph is here. This is about multiple
colours of rainbows.**

**Though it is a temporary event it leaves a great impact
on human mind. </p>**

<p> Best of Luck! </p>

Consider the second example from the above listed examples. You may notice the multiple blanks and carriage returns (where enter key is pressed and text appears on a new line) in the second paragraph. It is to be noted that such white space will be considered as a single space. When an HTML code is displayed into a browser, the text will automatically take up the full width of the screen. If you resize the browser window, the browser will wrap the text onto new lines. Such white space management allows developer of HTML code to indent the code and add extra spaces to maintain readability of the code.

Line Break:
 or

Unlike all other tags, the
 element does not have an opening and closing tags.
 is an abbreviated form of break. Such tags are known as empty tags. Advanced versions like XHTML use
 tag. Note that there should be a space between 'br' and '/'.

The
 just pushes the coming text into next line. In case you need multiple lines, simply use multiple
 tags. Examples demonstrating use of the
 are as follows:

**First example

**

Second example

Using
 in paragraph forces compulsory break into the content and disturbs text wrapping while presentation and resizing the browser window.

Preformatted Text

Many a times we want to display text with multiple white spaces and in multiple lines without wanting to be changed it by the browser. For that, we may embed the content into pre-formatted tag set using <pre> and </pre>. Any text between the opening <pre> tag and the closing </pre> tag will preserve the formatting of the given content. Example is as follows :

<pre>

This is first line.

This is second line.

This is third line.

</pre>

In a browser the text appears along with given indentations. Also try following examples.

**Bold : and **

This tag is used to display given content into bold letters. The example can be given as follows :

<p> This is the first paragraph. </p>

Here the "first" word is displayed in bold letters.

Underline : <u> and </u>

This tag is used to display given content with underlined letters. The example can be given as follows :

<p> This is the <u> first </u> paragraph. </p>

Here the "first" word is displayed in underlined manner.

Italics : <i> and </i>

This tag is used to display given content into italics letters. The example can be given as follows :

<p> This is the <i> first </i> paragraph. </p >

Here the "first" word is displayed in italics letters.

Strike Through : <s> and </s>

The content of an <s> or <strike> element is displayed with a strikethrough a thin line through the text. Here the 's' is an abbreviated form of 'strike'. Example of presentation of strikethrough content is given as follows.

<p> This is the <s> cancelled </s> paragraph. </p>

Here the "cancelled" word is displayed in strikethrough manner.

Type writer font: <tt> and </tt>

The content of a <tt> element is written in typewriter type of fonts, which is also identified as mono-spaced font (like that of a teletype machine). Example is as follows :

<p> This is the <tt> first </tt> paragraph. </p>

Here the "first" word is displayed in mono-spaced fonts.

There are some other elements which are described in table 1.2.

Elements	Description
<small> and </small>	The content is displayed one font size smaller than the rest of the text surrounding it.
<big> and </big>	The content is displayed one font size bigger than the rest of the text surrounding it.
^{and}	The content is displayed in superscript.
_{and}	The content is displayed in subscript.
<acronym> and </acronym>	It defines the content as an acronym.
<dfn> and </dfn>	It defines a special term.
<q> and </q>	It defines a quote.

Table 1.2 : Some other formatting tags

Anchor Tag

When text is displayed within an HTML document, besides the content and format specification, some extra information or reference to other entity is needed. Many times further explanation is also required. Set of such words or text that appears in different colour (generally blue and underlined) are called hyperlink. A hyperlink is created using an <a> element, where the 'a' stands for an anchor. Let us modify the file "p1.html" as shown in code listing 1.2.

```

<html>
  <head>
    <title> About Rainbow
  </title>
  </head>
  <body>
    <h1> RAINBOW </h1>
    <p> Rainbow consists of seven colours. These colours are Violet, Indigo, Blue, Green, Yellow, Orange and Red. They are also acronymed as VIBGYOR.
  </p>
    <p> Rainbow is caused by reflection of light in water droplets in the Earth's atmosphere, resulting in a spectrum of light appearing in the sky. It takes the form of a multi coloured arc.
  </p>
    <p>
      <a href= "p4.html" > Click here to visit Theory of Rainbow. </a>
    </p>
  </body>
</html>

```

Code Listing 1.2 : HTML code for showing use of hyperlink

Save the file as "p3.html". Figure 1.8 shows the output of this code when viewed in browser.



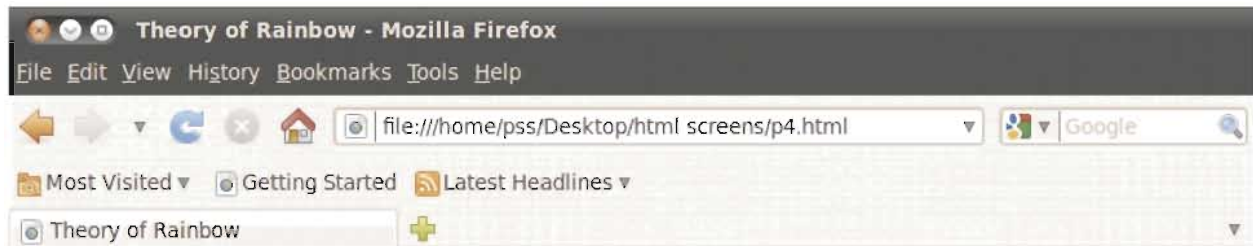
Figure 1.8 : Output of p3.html

When user clicks on the hyperlink, an intended file describing the 'Theory of Rainbow' must be opened. Let us create an HTML code for the file referred by the hyperlink as shown in code listing 1.3.

```
<html>
  <head>
    <title> Theory of Rainbow
  </title>
</head>
<body>
  <h1> How Rainbow Developed </h1>
  <p> Rainbow is caused by reflection of light in water droplets
    in the Earth's atmosphere, resulting in a spectrum of light
    appearing in the sky. It takes the form of multi coloured arc.
  </p>
</body>
</html>
```

Code Listing 1.3 : HTML code linked to hyperlink

Save the code shown in code listing 1.3 as "p4.html". When you click on the linked (anchor) text (Click here to visit Theory of Rainbow.) shown in figure 1.8. The contents of p4.html will be loaded in the browser, see figure 1.9.



How Rainbow Developed

Rainbow is caused by reflection of light in water droplets in the Earth's atmosphere, resulting in a spectrum of light appearing in the sky. It takes the form of a multi coloured arc.

Done

Figure 1.9 : Contents of p4.html

The hyperlink is a way to link two HTML documents by creating a hyper text in a document and giving reference of the other document to it. A website, which is a collection of many web pages, manages links through such hyperlink management. Here it is to be noted that, the web pages we have created are just presenting given content in a formatted way as we have specified them into the HTML code. Such web page contains only static (fixed) information, hence known as static web pages.

Absolute or Relative Address

Observe the line ` Click here to visit Theory of Rainbow. ` shown in code listing 1.2.

Instead of giving a full address such as `http://www.somedomain.com/p4.html`, we have given only the file name "p4.html". Giving just a file name will work only when you have the calling file (p3.html, also called parent file) and called file (p4.html, also called referred file) in the same directory. The location of the called file is relative to the calling file. Hence it is known as a relative address. While the complete address is known as an absolute address. If the referred file is located one directory above, we may prefix `../` to the filename.

Note :

If no path is provided, the browser will understand that the referred file is located in the same directory where the parent file is stored.

Attributes to the Tags

To specify more information along with tags, additional attribute accompany the tags. In other words, attributes tell more about the elements. Attributes always appear on the opening tags of the elements that carry them. An attribute is made up of two parts. The first part is a name and the second part is a value.

The name of an attribute indicates the property to be set. In case of <a> tag demonstrated in table 1.5 has name href. The value is a value to be set to the property. In case of the href, the value is p4.html (the reference to the link). The values should be in double quotation marks. Between the name and the value there should be an equal (=) sign. See next section for example showing how to define a tag with an attribute.

Align Attribute

The align attribute indicates whether the heading appears to the left, center, or right of the page. By default, the content is aligned to the left of the page. It can take three values as follows.

Left : The content is aligned to the left of the page.

Right : The content is aligned to the right of the page.

Center : The content is aligned to the center of the page.

Following are some examples demonstrating use of the align attributes :

<p align="right"> This content will be displayed in right aligned form </p>

<p align="center"> This content will be displayed in center position of the page </p>

There are some attributes which can appear along with every tag. Such attributes are called universal attributes. Align is such universal attribute. Being a universal attribute, the align attribute can also go with heading as shown below.

<h1 align="center"> Centered Heading </h1>

When content given in a paragraph is aligned, some spaces are automatically added for adjustment. The spaces which are inserted automatically are known as soft spaces. If users himself (manually) inserts some spaces, such hard spaces will be automatically deleted unless the content is written using <pre> and </pre> tag pairs.

Working with other Editor - SciTE

SciTE is a text editor based on a free source code editing component called Scintilla [<http://www.scintilla.org>]. It comes with complete source code and a license that permits use in any free project or commercial product. The interface of the SciTE is shown in figure 1.10.

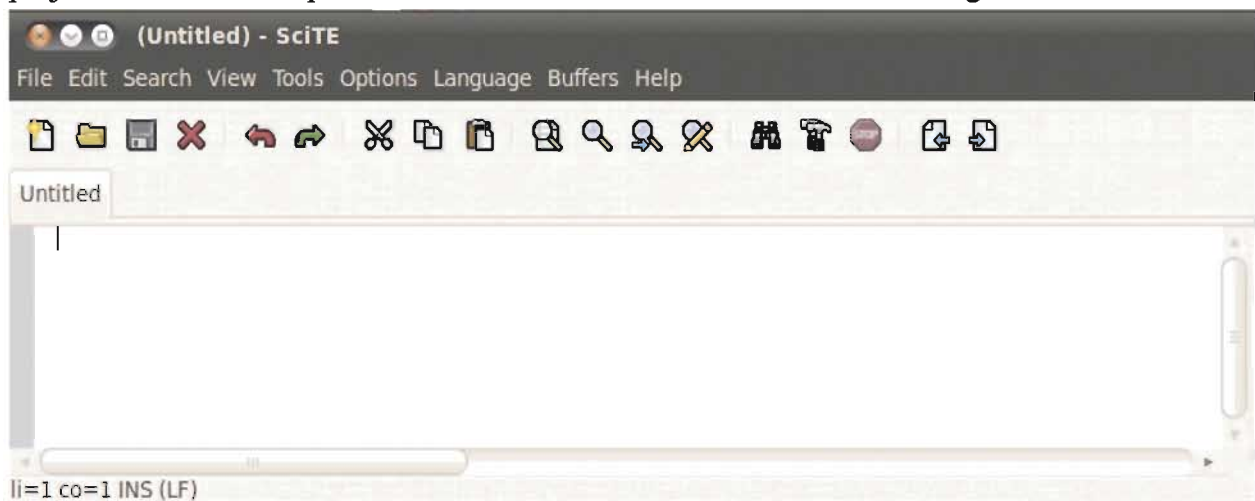


Figure 1.10 : Interface of the SciTE editor

The HTML code written above that prints information about rainbow (See example in figure 1.3) can be written in the SciTE editor as shown in figure 1.11.

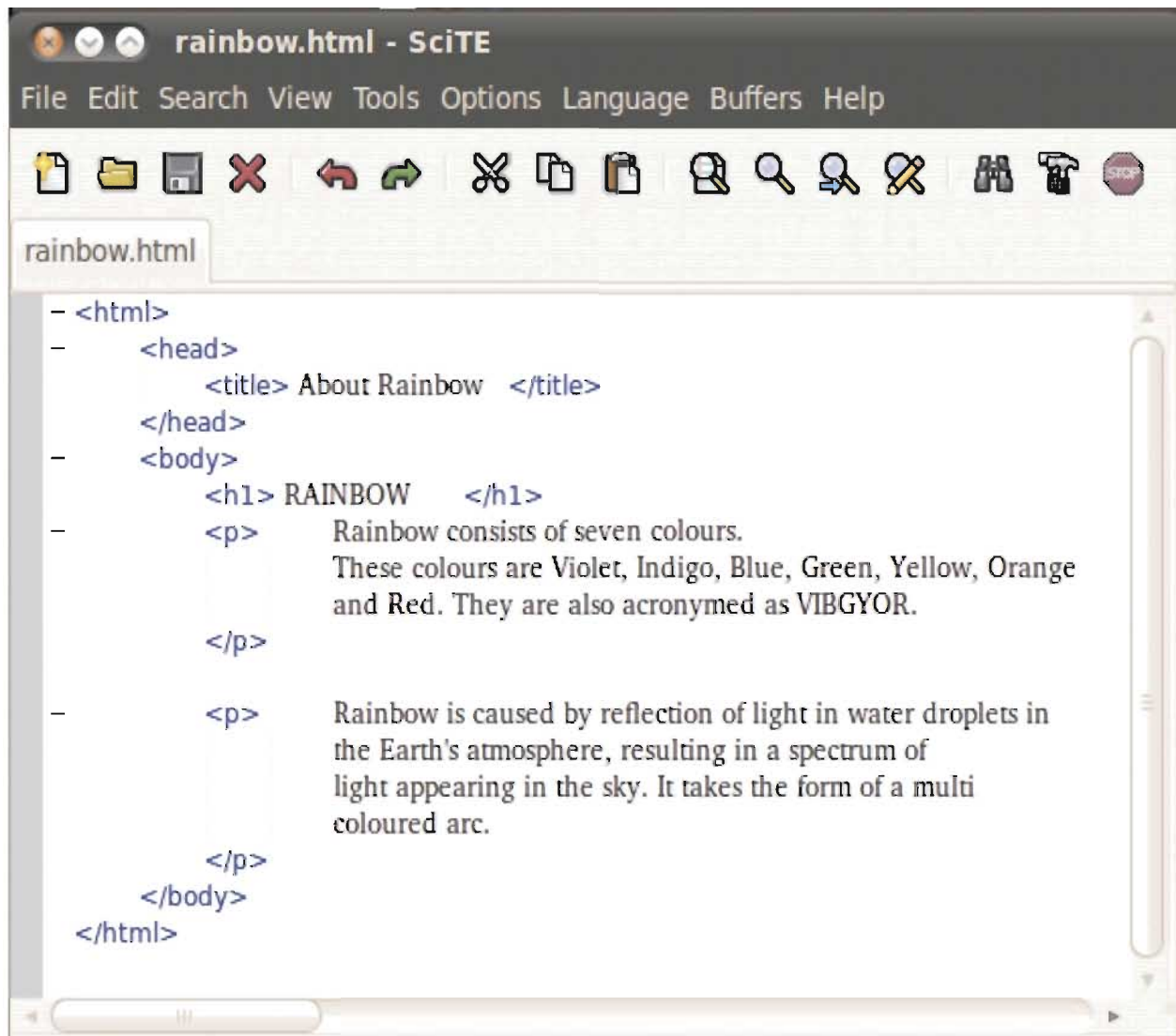


Figure 1.11 : HTML code in SciTE editor

Notice the menu bar items at the top of the screen shown in the figure 1.11. The menu items include basic facilities for file operations, editing facilities, searching, viewing, and other tools, options, etc. Next line beneath the menu items represents some icons for basic operations such as new, open, save, edit, search, etc. You may open any existing code in the SciTE editor as follows.

Step 1 : Locate the file you want to open.

Step 2 : Right click on it and choose Open With. You will see interface as shown in figure 1.12.

Step 3 : Select SciTE Text Editor. It will open the file in the SciTE editor. You may note the indicators for indentations and colour of tags shown by the SciTE. Having such indentations and tags in different colour separates the content from the tags and increases ease of reading the code.

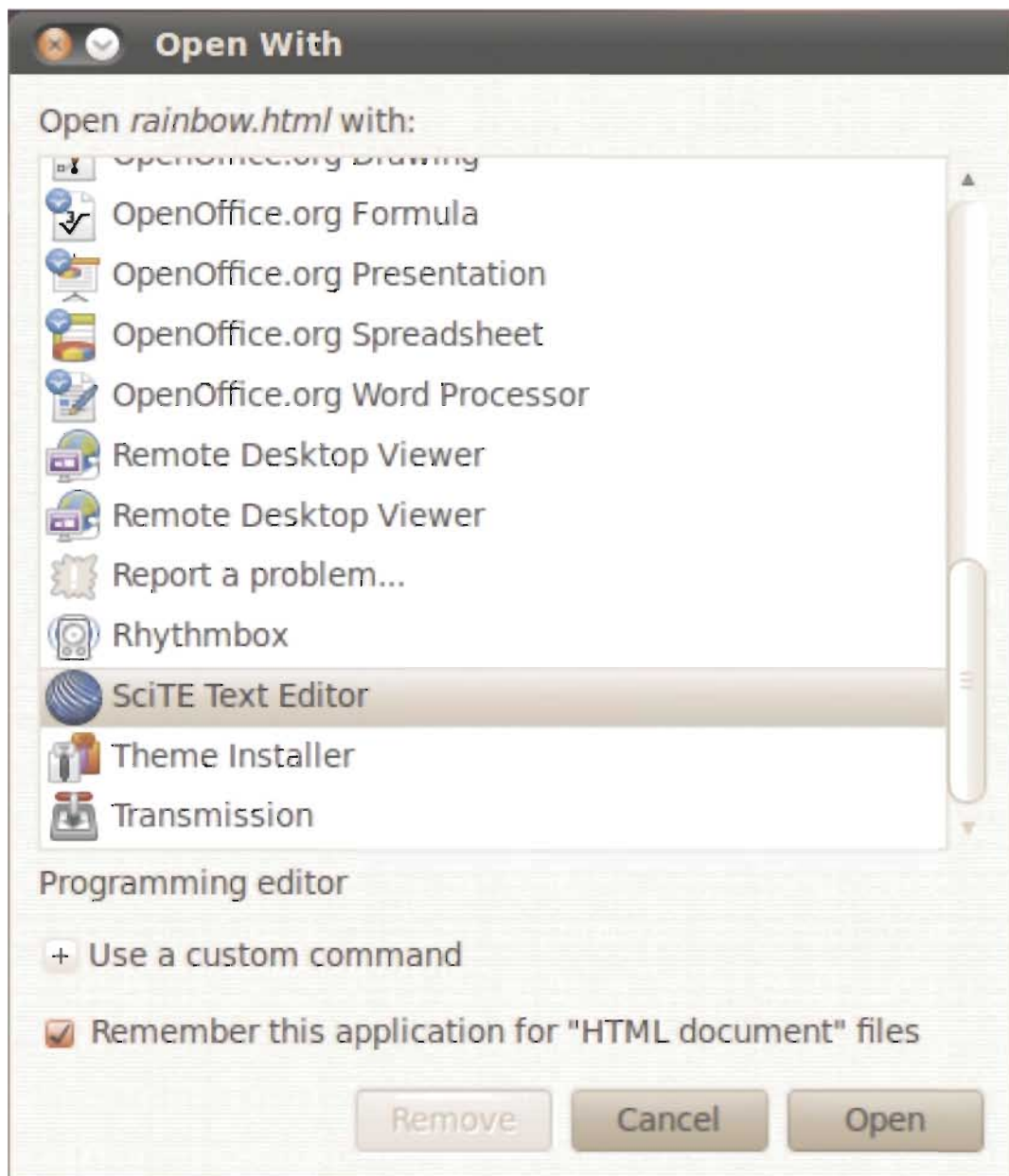


Figure 1.12 : The Open With dialog box

Step 4 : To view the html output in browser select Tools → Go or press F5.

Current Version of HTML

During the process of its evolution the HTML kept its focus on its purpose so that the composing and publishing content remains effective. With the progress of time, more and more functionalities were added to the HTML so that it can be useful and compatible with new browsers, new technologies and ever increasing developer demands. Latest version of the HTML is called HTML 5.0 and is becoming popular now.

Tags Covered in this Chapter

In this chapter we have discussed the tags mentioned in table 1.3.

Tag	Description
<code><a href> ... </code>	Anchors parent file to the referred file through hot text (link)
<code> ..</code>	Displays text in bold fonts.
<code><body>...</body></code>	Defines body of the HTML document. Appears within the <code><html></code> tag pair.
<code>
</code> or <code>
</code>	Defines line break. It is an empty singular tag.
<code><h1>...</h1></code>	Defines a first level heading.
<code><h2>...</h2></code>	Defines a second level heading.
<code><h3>...</h3></code>	Defines a third level heading.
<code><h4>...</h4></code>	Defines a fourth level heading.
<code><h5>...</h5></code>	Defines a fifth level heading.
<code><h6>...</h6></code>	Defines a sixth level heading.
<code><head>...</head></code>	Defines the head section of an HTML document. Appears within <code><html></code> tag pair.
<code><html>...</html></code>	Covers the entire HTML document.
<code><i> ..</i></code>	Displays text in italics fonts.
<code><p>...</p></code>	Defines a paragraph
<code><pre>...</pre></code>	Displays preformatted text.
<code><s> ..</s></code>	Displays text in strikethrough manner.
<code><title>...</title></code>	Defines title of the document. Appears within the <code><head></code> tag pair.
<code><tt> ..</tt></code>	Displays text in typewriter fonts.
<code><u> ..</u></code>	Displays text in underlined fonts.

Table 1.3 : HTML Tags Covered in Chapter 1

Summary

In this chapter we learnt how to design a web page using HTML. HTML is a fundamental utility that describes how the web content is composed, published and retrieved using web browsers. In this chapter, two important components of the structure of an HTML document such as head and body are described with some basic tags. Using the information you can create simple web pages. You may also link multiple web pages created using the anchor tag discussed in this chapter. Besides providing fundamental concepts of the HTML and history of the HTML; the chapter also provides information about editors such as gedit and SciTE to build HTML documents.

EXERCISE

1. Write a short note on history and evaluation of HTML.
2. Explain structure of HTML document by giving a simple example of an HTML document.
3. What is an HTML element? Give structure of an HTML element.
4. How to view an HTML document on your desktop?
5. List any three browsers.
6. Explain various HTML heading styles by giving examples.
7. Explain various HTML formatting tags by giving examples.
8. Explain preformatted text in HTML by giving an example.
9. Write a short note on anchor tag of HTML.
10. Define and explain absolute and relative address in HTML.
11. **Choose the correct option from the following :**
 - (1) To display the web content, which mark-up language is needed ?
(a) CML (b) HTML (c) NML (d) WML
 - (2) Which of the following is considered as a language for describing web page ?
(a) HTML (b) WML (c) NML (d) CML
 - (3) Which of the following is the full form of HTML ?
(a) Hot Text Manipulation Language (b) Hyper Text Manipulation Law
(c) Hyper Text Markup Language (d) Hidden Text Markup Language
 - (4) Which of the following is the full form of SGML ?
(a) Standardized General Markup Language
(b) System General Manipulation Law
(c) Standardized Genome Markup Law
(d) Standardized Gigabyte Markup Language
 - (5) Which of the following refers to an HTML element ?
(a) An opening tag, content and a closing tag (b) Angular brackets
(c) Content (d) Any of these
 - (6) Which of the following can be used to specify additional formatting along with an HTML element?
(a) Numbers (b) Attributes
(c) Comments (d) Contents
 - (7) Which of the following refers to a singular tags that do not require content ?
(a) Compete (b) Empty (c) Null (d) Void
 - (8) Which of the following attributes type can appear along with any tag ?
(a) Unique (b) Universal
(c) Trivial (d) Preliminary

- (9) Which type of information can be incorporated in an HTML document ?
- (a) Multimedia information (b) Text information
(c) Address and path of filename (d) All of these
- (10) Which of the following is an editor to edit an HTML document ?
- (a) SciTE (b) BriTE (c) LigHT (d) SpriTE

LABORATORY EXERCISE

1. Develop a web page that provides introductory information about your school. Give heading of this page as 'My School'. Use necessary formatting and presenting tags.
2. Develop a web page that introduces your class. Include information such as your class teacher, other course teachers and subjects you are learning. Give heading of this page as 'My Class'. Use necessary formatting and presenting tags.
3. Modify the web page created in question 1 of this exercise; to create a reference to another web page you have created in question 2 of this exercise. Set hot text in such a way that, when it is clicked, from the 'My School' page it will jump to the 'My Class' page.

