

Creating Your Own Web Pages

A Quick Reference to HTML



Introduction

HTML (hypertext markup language) is the "language" used to create World Wide Web documents or pages. It is not a language in the conventional sense of a computer language. Rather, it is a set of conventions for embedding tags or markup labels within a plain text file. These tags tell browsers (e.g., Netscape Navigator, Internet Explorer) how to display the information in the file and create links to other files. Any text editor or word processor that can save files in plain text (ASCII) format can be used to create HTML documents.

Today, there are many easy-to-use webpage editing programs (e.g., FrontPage, Dreamweaver) and even word processors such as Microsoft Word can save files in HTML format. So, why in the world would anyone want to know HTML itself? The answer is that a little knowledge can go a long way in some circumstances. Webpage editors and word processors sometimes yield poor results; if you understand HTML, you can often correct these problems. In addition, built-in features in web environments such as WebCT can be enhanced if you know how to use HTML to "dress up" what would otherwise be plain text. Some knowledge of HTML gives you more capability and a better understanding of how webpage editors and browsers work.

For more information about HTML, you can access Web documents at:

- <http://www.ncsa.uiuc.edu/General/Internet/WWW/HTMLPrimer.html>
(Beginner's Guide to HTML)
- <http://imbolc.ucc.ie/~pflynn/books/htmlcard.html> (Quick Reference Guide to HTML)
- http://dir.yahoo.com/Computers_and_Internet/Data_Formats/HTML/ (Yahoo)

or any of hundreds of other sites that have information about HTML and webpage creation.

HTML Document Basics

A basic Web document consists of a plain text file that contains embedded directions to the browser program, what are called tags, and often pointers to other files (e.g., graphics) as well as to other Web sites. Tags usually come in pairs and are of the form:

```
<TAG>Here is some text that gets formatted</TAG>
```

<TAG> is not a real HTML tag; it just illustrates the typical form. Each tag is enclosed in brackets (< >), and the slash (/) denotes an ending tag. Some tags define text styles, some tags create lists of items, and some create links to other information. In the case of a text formatting

tag, for example, the text between the tags has the formatting applied to it (e.g., `Some text` applies boldface tags to yield **Some text**). Tags themselves may be capitalized or not; browsers don't care. Some people capitalize all of their tags to make them easier to see; most people don't. Tags are capitalized in this handout just so that you can see them more easily.

A sample HTML document with the most common features is shown following. The tags, `<HTML>` and `</HTML>`, mark the beginning and end of the document. Documents have two main portions, the head and the body, each denoted with corresponding tags `<HEAD>` `</HEAD>` and `<BODY>` `</BODY>`. Often, the only really important thing in the head is the title of the webpage. The body is where the meat of the document can be found. Let's go on and see how some basic head and body elements are constructed.

Sample HTML document

```
<HTML>
<HEAD>
  <!-- A sample document -->
  <TITLE>Document title</title>
</HEAD>
<BODY>
  <H1>Top-level heading</H1>
  <P>First paragraph of text.</P>
  <UL> <!-- A bulleted list -->
    <LI>First list item
    <LI>Second list item, with a hypertext <a href="info.html">link<a> to another file.
  </UL>
  <H2>Second-level heading</H2>
  <IMG SRC="mypic.gif" ALIGN="bottom" ALT="An embedded picutre">
</BODY>
</HTML>
```

HEAD

As noted above, the Head portion of an HTML document doesn't usually contain a lot. However, you should be sure to give your page a title. To give the page a title, simply enclose whatever you want as a title between title tags as shown here:

```
<TITLE>My Home Page</TITLE>
```

Titles of less than 64 characters are recommended. When your page is loaded by a browser, the title is displayed at the top of the browser window.

Many web authors also include what are known as Meta tags in the Head. Meta tags are often used by search engines to locate relevant documents.

```
<META NAME=Author CONTENT="James D. Lehman">
```

```
<META NAME=Description CONTENT="EDCI 663 Home Page">
```

These tags, which are only some of the available Meta tags, define the author of the document as well as the content of the webpage. Proper use of Meta tags increases the chances that your webpages will be "noticed" by search engines.

BODY

Text Formatting

One common feature of Web documents is formatted text. With appropriate tags, you can do all sorts of different text formatting. For example, to write very large text -- major heading size -- you just use the following tags:

```
<H1>This is a first level heading</H1>
```

will produce something like:

This is a first level heading

There are six levels of headlines, H1 through H6. Each level gets progressively smaller.

If you want to set off just a regular paragraph of text, use the paragraph tags:

```
<P>This defines a paragraph.</P>
```

Paragraphs are separated from other text by a blank line on either side. If you want to start a new line at a specific location within a paragraph, use the
 tag to force a line break. Unlike most HTML tags, there is no paired ending tag for a break.

If you'd like to set off text like a paragraph but also indented, use the blockquote tags:

```
<BLOCKQUOTE>Blockquotes are indented.</BLOCKQUOTE>
```

You can also control the font elements of text with tags such as:

<I> . . . </I>	italics
 . . . 	boldface
<U> . . . </U>	underlined
<BIG> . . . </BIG>	big text
<SMALL> . . . </SMALL>	small text

Special characters are available using special symbols. These are not tags in the usual sense, but are interpreted by the browser to provide special symbols when you need them. A few of the many available ones are:

<	the less than symbol (<)
>	the greater than symbol (>)
&	the ampersand (&)
"	quote marks (")
 	a forced non-breaking blank space

The latter tag is needed because browsers, by default, insert only a single space between words regardless of how many spaces are typed by the author.

Text Alignment

The default text alignment is left justified, just like your favorite word processor. If you would like to center information on the screen, use the Center tag:

```
<CENTER>This text will be centered.</CENTER>
```

You can also use the `Align` attribute in association with one of the text formatting tags. For example, you can create a centered heading as follows:

```
<H1 ALIGN=center>This is a centered heading.</H1>
```

Legitimate values that can be used with `Align` are `left`, `right`, and `center`.

Lists

Many Web documents contain lists of items. Lists may be embedded within other lists. Web documents often have one of two basic kinds of lists: unordered or ordered. The individual items in a list are denoted with the `` or list item tag. An ending tag is not required for list items. List items are indented with respect to the surrounding text.

To create an unordered list use the tags, `` and ``, with list items:

```
<UL>
  <LI>item 1
  <LI>item 2
  ...
</UL>
```

List items in an unordered list are bulleted.

To create an ordered list use the tags, `` and ``, with list items:

```
<OL>
  <LI>item 1
  <LI>item 2
  ...
</OL>
```

In an ordered list, list items are numbered consecutively.

Images

To add an inline image, use an `` or image tag that points to a graphic file that is stored separately on your Web server. On the WWW, two graphic formats are universally accepted: GIF (a graphic interchange format originally developed by CompuServe) and JPEG or JPG (a file format good for storing high color images such as photos). If you have a graphic that you would like to incorporate in a Web page that is in some other format (e.g., Macintosh PICT or Windows BMP), use conversion software to convert the image to GIF or JPEG format first.

To define a graphic image, the basic tag looks like this:

```
<IMG SRC="url">
```

where `SRC` stands for the source or the image, and `URL` is the location of the graphic file. If all your Web files are kept in the same folder on the server, you can just give the filename (e.g., `graphic.gif`); otherwise, use a full URL (e.g., `http://www.soe.purdue.edu/~lehman/graphic.gif`).

A couple of optional attributes are often included as part of the IMG tag. ALIGN can be used to define where text following the tag appears relative to the graphic (top, middle, or bottom) or to force the graphic to one side of the page (left or right). ALT is used to provide a text label that is used if the browser has graphics loading turned off or if a screen reader is being used by a visually impaired individual. So, a more complete example might look like:

```
<IMG SRC="graphic.gif" ALIGN=middle ALT="A neat picture.">
```

Hypertext Links

One of the great things about documents on the Web is that they link to other documents. Links are accomplished using the <A> or anchor tag. The basic format of an anchor tag is like this:

```
<A HREF="url">the text that signifies the link</A>
```

The first part of the expression indicates what URL, or Web location, will be linked to. The text in the middle is what shows up as an active link in the browser (usually denoted by the color blue and underlined). So, a link to Jim Lehman's home page would be denoted by:

```
<A HREF="http://www.edci.purdue.edu/lehman">Jim Lehman's home page</A>
```

The text "Jim Lehman's home page" would appear underlined and blue in the browser. Clicking on that text with the mouse would execute the link to the actual website. Links such as this can be embedded in the middle of sentences to make the links appear as natural parts of what is written.

URLs may begin with any of the following identifiers to specify a resource for a hypertext link:

http:	Web site
gopher:	Gopher site
ftp:	FTP site
telnet:	Telnet site
mailto:	E-mail address
news:	Newsgroup

Note mailto: and news: use no double slash following when writing the URL; all of the others do.

Other Goodies

Comments

If you want to add comments, notes to yourself in your Web documents, set them off with the following tags:

```
<!-- a comment -->
```

Horizontal Rules

You can add horizontal lines, or what are called horizontal rules, to act as visual dividers with the <HR> tag. This tag has no ending counterpart. Just put one wherever you want a horizontal line to appear.

Colors

Colors can be added in a couple of ways. You can define a background color for an entire page as part of the Body tag as follows:

```
<BODY BGCOLOR="aqua">
```

Certain basic colors are recognized by name (e.g., blue, green, red, aqua, black). In addition, you can define colors by number (e.g., #000000) where each two hexadecimal digits correspond to hues of red, green, and blue respectively. Specifying colors this way can be pretty hit or miss without a table of values. A few basic color numbers are: #000000 - black, #FFFFFF - white, #FF0000 - red, #00FF00 - green, #0000FF - blue, and #3299CC - sky blue.

You can also have an image serve as the background of a page by using:

```
<BODY BACKGROUND="url">
```

where the URL refers to the location of a graphics file. The browser will automatically tile the image to turn it into the page background. Avoid using strong images as page backgrounds; they can easily overwhelm your text.

You can also add color to individual passages of text by using the Font tag with a color as in:

```
<FONT COLOR="red">This text will be red.</FONT>
```

Tables

Finally, a great way to organize information is with the Table tag and its corresponding tags. The basic definition of a table looks like this:

```
<TABLE BORDER=2>
<CAPTION ALIGN=top>Table caption</CAPTION>
<TR><TH>Heading 1<TH>Heading 2
<TR><TD>Data set<TD>Data set
<TR><TD>Data set<TD>Data set
...
</TABLE>
```

The Table tags define the beginning and end of a table. Border defines the width of the border lines around each cell of the table. The Caption tag sets the caption for the table. <TR> defines a row in the table. <TH> defines a heading for a table column. <TD> defines an element of data that occupies one cell in the table. It is the author's responsibility to ensure that the proper number of headings and data entries are defined to fill out the table correctly, i.e. if you define four headings then there should be four sets of data in every row. Tables can be useful for displaying organized information such as a class schedule.

Putting Your Page Up for the World to See -- SOE Faculty

After you have created a document with tags that will become a World Wide Web page, save the document in plain text format. Your main Web document, the one that will first be accessed when someone comes to your Web site, should be saved with a special name:

```
welcome.html or welcome.htm or index.htm
```

Although not absolutely necessary, it is often safest to adhere to file naming conventions for MS-DOS (up to 8 characters, no spaces, and an extension) with the extension .html or .htm to identify the files as HTML files. Names like intro564.htm and syll564.htm are useful to identify

webpages associated with a course. While capitalization is not important on your personal computer, it often is important on web servers. So, be consistent (e.g., stick with no capital letters in any filenames).

Next, you have to have a place to put your Web documents so that people browsing the WWW can find them. You have to place your files on a server computer that is connected to the Internet. The School of Education maintains a World Wide Web server. You can place your information on that server for others to access.

In a nutshell, here is the process. Before you get started, obtain the ID and password needed to log in to the SOE WWW server. (**Note:** please do not share the ID and password information with anyone; it could leave the SOE server vulnerable to outside access. Bob Evans periodically changes the information to maintain security; check with him for updates.)

Use FTP software to connect to the server and transfer your files to it:

- (1) Use FTP software to connect to the server. Use your ID and password to log in.
- (2) If necessary, navigate to your own folder on the server.
- (4) Now, transfer your prepared files from your PC to your folder on the server. **Important:** if your FTP software does not automatically do so, be sure to select text transfer for any HTML files, but be sure to select binary transfer for any GIF or JPG image files. You must transfer the HTML files and any associated image files for the pages to work properly. Remember, your main HTML file should be named welcome.html or welcome.htm or index.htm.
- (5) That's it. Your page(s) will be accessible via the appropriate server, such as:
<http://www.edci.purdue.edu/lastname>
- (6) Any time you want to add or update information, repeat these steps to use FTP software to upload your files to the server. Keep copies of your HTML files on your personal computer. When you make an update, simply resave the file (in text only format), FTP it back into your folder on the server, and the new version will immediately be accessible on the WWW.

Epilogue

A short example of a Web page follows. Study it and use View Page Source in Navigator or View Source in Explorer to see how other pages are constructed. That's the best way to learn. There is much more to HTML. There are many more advanced features that you may want to try as you become more comfortable. But, what you have here is enough to get started making information available to others on the Internet.

A Simple Example -- EDCI 564 Home Page

```
<HTML>

<HEAD>
<TITLE>EDCI 564 Home Page</TITLE>
</HEAD>

<BODY>
<IMG ALIGN=top SRC="http://www.edci.purdue.edu/lehman/edci564/564pict.gif"
ALT="EDCI 564, Integration and Management of Computer in Education, Purdue University">

<CENTER>
<H1>EDCI 564</H1>
<H2>Spring, 2001</H2>
</CENTER>

<P>Welcome to the EDCI 564 home page. This WWW page was established to give students
and interested others access to information for and about EDCI 564, Integration and Management
of Computers in Education, at Purdue University.</P>

<CENTER>
<HR>
<H3>Links to Additional Course Information</H3>
<UL>
<LI><A HREF="syll564.htm">Course Syllabus</A>
<LI><A HREF="http://www.ed.gov/Technology/techno.html">U.S. Department of Education
technology reports</A>
</UL>

<HR>
<H3>Purdue Links</H3>
<UL>
<LI><A HREF="http://www.soe.purdue.edu/">Purdue University School of Education</A>
<LI><A HREF="http://www.purdue.edu/">Purdue University Home Page</A>
</UL>

<HR>
<I>Created by Jim Lehman. Maintained by Jim Lehman. Last update: January 8, 2001.</I>

</CENTER>
</BODY>
</HTML>
```