

LESSON 3

Adding Text and More



In this lesson, you will learn how to use HTML to add text and headings in your Web pages. You'll also learn how to add mathematical notations, information about your Web page, and special characters (such as ampersands).

Paragraphs

You might not realize it, but you already learned how to create an HTML paragraph in Lesson 2, “Creating Your First Page.” In HTML, a paragraph is created whenever you insert text between the `<p>` tags. Look at the code from Lesson 2 again:

```
<!DOCTYPE html
    PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-
    transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
    xml:lang="en" lang="en">
<head>
<title>My XHTML Page</title>
</head>
<body>
<p>This is my first XHTML page.</p>
</body>
</html>
```

Web browsers see that you want text and they display it. Web browsers don't pay any attention to how many blank lines you put in your text; they only pay attention to the HTML tags. In the following HTML code, you see several lines of text and even a blank line, but the browser only recognizes paragraphs surrounded by the `<p>` and `</p>` tags (or paragraph tags). The `<p>` tag tells the browser to add a blank line before displaying any text that follows it, as shown in Figure 3.1.

```
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-
  transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
  xml:lang="en" lang="en">

<head>
<title>Typing Paragraphs in HTML</title>
</head>
<body>
<p>This is the first line.

But is this the second?</p>
<p>No, this is.</p>
</body>
</html>
```



FIGURE 3.1 The browser ignores the blank line that I inserted and puts the line break before the `<p>` tag instead.

Web browsers do something else with paragraph text that you should be aware of: They wrap the text at the end of the browser window. In other words, when the text in your Web page reaches the edge of the browser window, it automatically continues on the next line regardless of where the `<p>` is located. The `<p>` tag always adds a blank line, but you might not always want a blank line between lines of text. Sometimes you just want your text to appear on the next line (such as the lines of an address or a poem). You can use a new tag for this—the line break, or `
` tag, shown in Figure 3.2.

This new tag forces the browser to move any text following the tag to the next line of the browser, without adding a blank line in between. Figure 3.3 shows how the browser uses these two tags to format your text.



```

twinkle.htm - Notepad
File Edit Format Help
<!DOCTYPE html
PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
<head>
<title>Twinkle Twinkle</title>
</head>
<body>
<p>Twinkle, twinkle, litte star,<br />
how I wonder what you are.</p>
<p>Up above the world so high,<br />
like a diamond in the sky.</p>
<p>Twinkle, twinkle, litte star,<br />
how I wonder what you are.</p>
</body>
</html>

```

FIGURE 3.2 The `<p>` and `
` tags help to separate your text into lines and paragraphs.

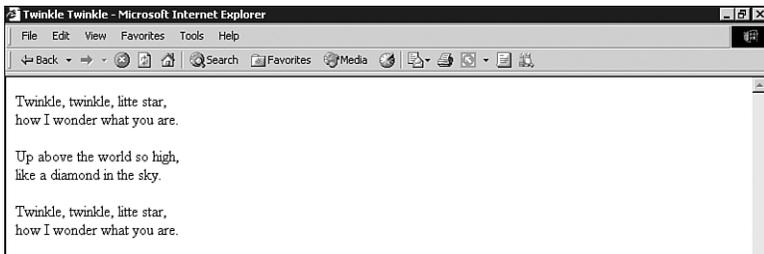


FIGURE 3.3 The browser inserts line breaks and blank paragraph separators only where you place the correct HTML tags.

Text Emphasis

So far you've learned how to add text, but here you will learn how to format it. You will occasionally want to add emphasis to your text to make it stand out. HTML enables you to quickly apply some standard formats, such as boldface and italic, using a predefined set of tags. All these tags occur in pairs (corresponding opening and closing tags) and must surround the text that they are emphasizing. Use the code that follows in your own Web page to see how each of these tags appears in the browser.

```

<!DOCTYPE html
PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-
transitional.dtd">

```

```
<html xmlns="http://www.w3.org/1999/xhtml"
      xml:lang="en" lang="en">
<head>
<title>Emphasizing Text</title>
</head>
<body>
<p>Make your text one size larger with the
  <big>big tag.</big></p>
<p>Try the <b>bold tag</b> or the <strong>strong tag</strong>
  to make an impact.</p>
<p>The <i>italics tag</i> and the <em>emphasis tag</em>
  create a different impact.</p>
<p>Use the <tt>teletype tag</tt> to imitate a typewriter.</p>
<p>Make your text one size smaller with the
  <small>small tag.</small></p>
</body>
</html>
```



Caution Other formatting tags exist in HTML, but their use is discouraged in HTML and deprecated in XHTML in favor of style sheets. The World Wide Web Consortium (W3C) has determined that HTML should be used to identify types of information (text, headings, tables, and so on), but should not be used to format that information.



Deprecated Some older HTML tags, specifically related to formatting, have been replaced by the formatting capabilities of style sheets.

Style Sheets Web developers use style sheets to specify formatting instructions for a single document or a group of documents.

You'll learn how to create style sheets in Lesson 5, "Adding Your Own Style." Throughout the book, however, you'll see how styles can enhance your Web pages.

Headings

Separating your text into paragraphs isn't the only way to format your Web pages. HTML enables you to add six different heading tags to your pages by using the tags `<h1>`–`<h6>`. These tags are very simple to use. Look at the following line of code:

```
<h1>This is Heading 1</h1>
```

The closing heading tags also create an automatic paragraph break. In other words, all headings automatically include a blank line to separate them from the text. Heading 1, the `<h1>` tag, has the largest font of the heading tags and Heading 6, the `<h6>` tag, has the smallest. In fact, you usually only see Web page authors use the `<h1>`–`<h3>` tags because the remaining tags, `<h4>`–`<h6>`, are actually smaller than normal text. Figure 3.4 shows a sample of all the heading tags compared to normal text.



FIGURE 3.4 Notice that HTML's Heading 4 is the same size as normal text, but Headings 5 and 6 are actually smaller.



Tip Unless you or the people viewing your pages have adjusted the browser's default settings, normal HTML body text appears in 12 point Times New Roman font on most computer systems.

Special Characters

You might find that you sometimes need to use symbols on your Web pages. Symbols (such as +, -, %, and &) are used frequently in our everyday writing, so it's easy to understand that they would appear on a Web page as well. Unfortunately, not all Web browsers display these symbols correctly. HTML uses a little computer shorthand, either using a numerical code or a text code (called an *entity character reference*) to tell the browser how to interpret these symbols. Table 3.1 shows some of the most frequently used codes.

TABLE 3.1 Special Character Codes

Char	Code	Description
&	&	Ampersand
<	<	Less than
>	>	Greater than
©	©	Copyright
®	®	Registered trademark
±	&plusmin;	Plus or minus
²	²	Superscript 2
³	³	Superscript 3
´	´	Acute accent
`	`	Grave accent
#	#	Number
%	%	Percent

Appendix C, “Special Characters,” contains a more complete list of the characters supported by HTML. You can see how many of these symbols are easy to understand (for example, & for the ampersand and > for the greater than symbol). Some of the characters, such as number and

percent, require that you memorize numbered codes. Yuck. The best thing you can do is to make sure that you preview your Web pages in a variety of browsers before publishing them.



Tip Here's a special character that you should remember: ` `. The symbol stands for *nonbreaking space* and is used to insert a space inside an HTML document. Because HTML ignores extra spaces between words and tags, you'll need to have a way of including extra spaces. You can do that with the ` ` character.

Math and Science Notations

Although HTML was first designed and used by scientists, it has yet to support mathematical and scientific notation with any degree of complexity. HTML does give you two tags to help write simple equations. Together with the codes for special characters, the `<sub>` (subscript) and `<sup>` (superscript) tags go a long way toward creating equations, as shown in Table 3.2.

TABLE 3.2 `<sup>` and `<sub>` Tags

You Type	The Browser Displays
<code>A<sup>2</sup> + B<sup>2</sup></code> <code>= C<sup>2</sup></code>	$A^2 + B^2 = C^2$
<code>CO<sub>2</sub> = Carbon Dioxide</code>	$CO_2 = \text{Carbon Dioxide}$

If you are looking to write more complex equations, you need to be a little more creative. The obvious answer is to write your equation in the program that you usually use, and then use a graphics program to turn it into an image. You can insert that image into any HTML page, as you've already learned. That works, but the solution is limited. Because the equation is graphical, you are not able to index or search for text within the equation. That's a big drawback, but so is the fact that images slow down

your page's load time and the fact that your equation cannot be viewed by nongraphical Web browsers.



Tip Some commercial products are available to help you notate mathematical expressions. You can see a list of them on the W3C Web site (www.w3.org/Math/).

English Isn't the Only Language

You can use HTML even if you don't write in English. URLs, hyperlinks, HTML tags, and document formatting elements are language neutral, but text requires a specification all its own. If you write in standard U.S. English, you don't need to make any changes to the way you create your HTML documents. If you are writing text in any other language, however, you should specify the language for the browser. The following HTML samples show the designations for German and French.

```
<html xmlns="http://www.w3.org/1999/xhtml"
      xml:lang="de" lang="de">
```

and

```
<html xmlns="http://www.w3.org/1999/xhtml"
      xml:lang="fr" lang="fr">
```

The language attributes (`xml:lang` and `lang`) support the same values as ISO, the International Standards Organization. You can see the full list of supported languages and their codes at www.loc.gov/standards/iso639-2/langcodes.html.



Tip Why is language important? Browsers do not recognize the language you type unless you use the `lang` attribute. Some search engines use the `lang` attribute to return only pages written in a specific language. Speech synthesizers use this information to aid in pronunciation. Even some spelling checkers can use the information to recognize misspellings.

Mixing Languages in a Single Page

Although the preceding example shows the `lang` attribute used as part of the `<html>` tag at the top of your document, it's possible that you would want to include text of one language within a document of another language—for example, including a paragraph in French within a document in English. You can assign the `lang` attribute to the `<p>` tag to solve this problem. Look at the following sample:

```
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-
  transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
  xml:lang="en" lang="en">
<head>
<title>Multi-Language Document</title>
</head>
<body>
<p>Put your English text here.</p>
<p lang="fr">Mettez votre texte français ici.</p>
<p lang="en">Put the rest of your English text here.</p>
</body>
</html>
```

Meta Tags

Finally, you get to do something with the `<head>` tag. So far, you've only seen the `<title>` tag used to give information about the document, but you can do a lot more with the `<head>` tag. What's more, aside from the `<title>` tag, information in the `<head>` tag doesn't usually appear in your document. You can use the meta information tag (`<meta>`) to identify the page's author, keywords used for searching, or a brief description to appear in search results. You also can use the `<meta>` tag to give commands to the browser. You can use as many `<meta>` tags as you like in your page. You'll learn how in the sections that follow.

Improved Searching

Search engines (as you'll find in Lesson 17, "Planning for the Future") add the content of your Web pages to their indexes. When a potential

visitor enters a search phrase, the search engine checks its index to find that word and returns any pages that include that word. It works great. But, what if you were a realtor and you worked hard at creating a Web page that included the words *houses*, *housing*, *sale*, and *buy*; but didn't include the phrase *real estate*? If that was the phrase your visitor was looking for, they would never find your page. You can use the `<meta>` tag to include product names, geographic locations, industry terms, and synonyms that people might be searching for. There are three `<meta>` tags that work to help improve your chances of being found by a search engine:

- **Keywords** Keywords are words that you feel people might use to search for your Web page, or synonyms for words that appear in your document.
- **Description** This is usually a paragraph of information about your page. Some search engines use the information in this tag to summarize your page, but other search engines use the first few lines of text in your actual document.
- **Author** This is your opportunity to shine. Just in case someone is searching for your name, they will find your page if you enter that information into the `<meta>` tag.

Meta information for search engines comes in pairs: name and contents. The following HTML code includes meta information pairs for each of the preceding `<meta>` tags.

```
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-
  transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
  xml:lang="en" lang="en">
<head>
<title>Your HTML Page</title>
<meta name="keywords" contents="words that people might
  use to search for your page." />
<meta name="description" contents="a brief paragraph describing
  your document." />
<meta name="author" contents="your name" />
</head>
<body>
```

```
<p>Insert your text here.</p>
</body>
</html>
```



Caution Remember, `<meta>` tags only appear in the `<head>` section of an HTML file.

Refresh and Redirect

There might be times when you want to replace one page with another or want to redirect a link. You might, for example, choose to include a *splash page* on your Web site. You can use the meta information to force the page to change within a given time span using the sample code that follows:

```
<meta http-equiv="refresh" content="time in seconds,
                                URL of the new page" />
```



Splash Page The introductory page used by some Web page authors to show flashy graphics or a product logo before continuing to the rest of the site's contents.

If you have a page that you update several times a day and you want to make sure that people always see the most recent version, you can enter the page's own URL in the refresh tag. When the browser sees the refresh tag, it presents the requested URL in the specified time.

```
<meta http-equiv="refresh" content="time in seconds,
                                URL for this page" />
```



Caution Because not all Web browsers support this attribute, authors should include some content on the splash page that enables users to move to the next page on their own.

Expiration Dates

If you have a page that you change frequently, you can specify an expiration date in the `<meta>` tag to ensure that the Web browser looks for a newer version (rather than displaying an older version, which might still be stored in the browser's memory). Look at the example that follows:

```
<meta http-equiv="expires"
      contents="Wed, 04 December 2006 00:00:00 GMT" />
```

When you enter the URL for this page in your browser, it checks its history files to see whether a copy is stored there. If so, it checks the meta information to see whether this page is still valid. If the expiration date has passed, the browser looks to the Web for a more recent copy before displaying the page.

Table 3.3 reminds you of the formatting tags you learned in this lesson.

TABLE 3.3 HTML Tags Used in This Lesson

HTML Tag	Closing	Description of Use
<code></code>	<code></code>	Text appears boldface.
<code><big></code>	<code></big></code>	Text appears one size larger than normal.
<code>
</code>		Line break. Forces text to the next line.
<code></code>	<code></code>	Text appears emphasized (italic). Usually the same as <code><i></code> .
<code><h1></code>	<code></h1></code>	A first-level heading.
<code><h2></code>	<code></h2></code>	A second-level heading.
<code><h3></code>	<code></h3></code>	A third-level heading.
<code><h4></code>	<code></h4></code>	A fourth-level heading. Rarely used.
<code><h5></code>	<code></h5></code>	A fifth-level heading. Rarely used.
<code><h6></code>	<code></h6></code>	A sixth-level heading. Rarely used.
<code><i></code>	<code></i></code>	Text appears emphasized (italic).
<code><meta /></code>		Identifies information about the document.

HTML Tag	Closing	Description of Use
<code><p></code>	<code></p></code>	Paragraph break. Forces a blank line.
<code><small></code>	<code></small></code>	Text appears one size smaller than normal.
<code></code>	<code></code>	Text appears boldface. Same as <code></code> .
<code><sub></code>	<code></sub></code>	Text appears in subscript.
<code><sup></code>	<code></sup></code>	Text appears in superscript.
<code><tt></code>	<code></tt></code>	Text appears monospaced, as if typed.

In this lesson, you've learned:

- The `<p>` tag, or paragraph tag, tells the browser to add a blank line before it displays any text that follows. The `
` tag moves your text to the next line without adding a blank line.
- HTML enables you to add emphasis to your text with several predefined formatting tags.
- Symbols such as `+`, `-`, and `%` require a little computer shorthand to tell the browser how to interpret these symbols. This shorthand begins with an ampersand (`&`) and ends with a semicolon (`;`). A more complete list can be found in Appendix C.
- You can add other languages to your HTML documents by using the `lang` attribute on the `<html>` tag.
- Meta information for search engines comes in pairs: name and contents, and the `<meta />` tags always appear between the `<head>` tags.