Chapter 5 Text and Lists

In This Chapter

- ▶ Working with basic blocks of text
- Manipulating text blocks
- Creating bulleted, numbered, and definition lists

TML documents consist of text, images, multimedia files, links, and other pieces of content that you bring together into one page by using markup elements and attributes. You use blocks of text to create such document elements as headings, paragraphs, and lists. The first step in creating a solid HTML document is laying a firm foundation that establishes the document's structure.

Formatting Text

Here's a super-ultra-technical definition of a *block of text*: some chunk of content that wraps from one line to another inside an HTML element.

Your HTML page is a giant collection of blocks of text:

- Every bit of content on your Web page must be part of some block element.
- Every block element sits within the <body> element on your page.

HTML recognizes several kinds of text blocks that you can use in your document, including (but not limited to)

- ✓ Paragraphs
- Headings
- Block quotes
- 🖊 Lists
- 🖊 Tables
- 🖊 Forms

Inline elements versus text blocks

The difference between inline elements and a block of text is important. HTML elements in this chapter describe blocks of text. An *inline element* is a word or string of words *inside* a block element (for example, text emphasis elements such as or). Inline elements must be

nested within a block element; otherwise, your HTML document isn't syntactically correct.

Inline elements, such as linking and formatting elements, are designed to link from or change the appearance of a few words or lines of content found inside those blocks.

Paragraphs

Paragraphs are used more often in Web pages than any other kind of text block.



HTML browsers don't recognize the hard returns that you enter when you create your page inside an editor. You must use a $\langle p \rangle$ element to tell the browser to separate the contained block of text as a paragraph.

Formatting

To create a paragraph, follow these steps:

- 1. Add in the body of the document.
- 2. Type the content of the paragraph.
- 3. Add to close that paragraph.

Here's what it looks like:

```
<!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">
<head>
    <meta http-equiv="Content-Type" content="text/html; charset=ISO-8859-1" />
    <title>All About Blocks</title>
    </head>
</body>

This is a paragraph. It's a very simple structure that you will use
    time and again in your Web pages.

This is another paragraph. What could be simpler to create?
```

This HTML page includes two paragraphs, each marked with a separate $\langle p \rangle$ element. Most Web browsers add a line break and full line of white space after every paragraph on your page, as shown in Figure 5-1.





Some people don't use the closing $\langle /p \rangle$ tag when they create paragraphs. Although some browsers let you get away with this, leaving out the closing tag

- Doesn't follow correct syntax
- \checkmark Causes problems with style sheets
- ✓ Can cause a page to appear inconsistently from browser to browser

You can control the formatting (color, style, size, and alignment) of your paragraph by using Cascading Style Sheets (CSS), which we cover in Chapters 8 and 9.

Alignment

By default, the paragraph aligns to the left. You can use the align attribute with a value of center, right, or justify to override that default and control the alignment for any paragraph.

```
This paragraph is centered.
This paragraph is right-justified.
This paragraph is double-justified.
```

Figure 5-2 shows how a Web browser aligns each paragraph according to the value of the align attribute.



The align attribute has been deprecated (rendered obsolete) in favor of using CSS (see Chapter 8).

Headings

Headings break a document into sections. This book uses headings and subheadings to divide every chapter into sections, and you can do the same with your Web page. Headings can

- Create an organizational structure
- Break up the visual appearance of the page
- \checkmark Give visual clues about how the pieces of content are grouped

HTML includes six elements to help you define six different heading levels in your documents:

- ✓ <h1> is the most prominent heading (Heading 1)
- ✓ <h6> is the least prominent heading (Heading 6)



Follow heading order from highest to lowest as you use HTML heading levels. That is, don't use a second-level heading until you've used a first-level heading, don't use a third-level heading until you've used a second, and so on. If you want to change how headings appear in a browser, Chapter 8 and Chapter 9 show you how to use style sheets.



Formatting

To create a heading, follow these steps:

- 1. Add <hn> in the body of your document.
- 2. Type the content for the heading.
- 3. Add </h*n*>.

Browser displays

Every browser has a different way of displaying heading levels, and we cover that in the following two sections.

Graphical browsers

Most graphical browsers use a distinctive size and typeface for headings:

- First-level headings (<h1>) are the largest (usually two or three font sizes larger than the default text size for paragraphs).
- ✓ Sixth-level headings (<h6>) are the smallest and may be two or three font sizes *smaller* than the default paragraph text.

The following excerpt of HTML markup shows all six headings at work:

```
<!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">
  <head>
   <meta http-equiv="Content-Type" content="text/html; charset=ISO-8859-1" />
   <title>All About Blocks</title>
  </head>
 <body>
   <h1>First-level heading</h1>
   <h2>Second-level heading</h2>
   <h3>Third-level heading</h3>
   <h4>Fourth-level heading</h4>
   <h5>Fifth-level heading</h5>
   <h6>Sixth-level heading</h6>
  </body>
</html>
```

Figure 5-3 shows this HTML page as rendered in a browser.

You can use CSS to format such heading aspects as color, size, line height, and alignment.



By default, most browsers use Times New Roman fonts for all headings. The font size decreases as heading level increases. (Default sizes for first-through sixth-level headings are, respectively, 24, 18, 14, 12, 10, and 8.) You can override any of this formatting by using CSS.

Part II: Formatting Web Pages with (X)HTML



Text browsers

Text-only browsers use different heading conventions than graphical browsers because text-only browsers display all content using a single size and font.

Controlling Text Blocks

Blocks of text are the foundation for your page. You can break those blocks to better guide readers through your content.

Block quotes

A *block quote* is a long quotation or excerpt from a printed source that you set apart on your page. You use the <blockquote> element to identify block quotes:

```
<!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">
    </html xmlns="http://www.w3.org/1999/xhtml">
    </html xmlns="http://www.w3.org/1999/xhtml">
    </http://www.w3.org/1999/xhtml">
    </http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1999/xhtml</http://www.w3.org/1
```

```
<body>
    <br/>
        <bdy>
        <br/>
        <br/>
```

Most Web browsers display block-quote content with a slight left indent, as shown in Figure 5-4.

Preformatted text

Ordinarily, HTML ignores white space inside documents. A browser won't display a block element's

- Hard returns
- ✓ Line breaks
- 🖊 🖊 Large white spaces

The following markup includes several hard returns, line breaks, and a lot of space characters. Figure 5-5 shows that the Web browser ignores all of this.

```
This is a paragraph
with a lot of white space
thrown in for fun (and as a test of course).
```

The preformatted text element () instructs browsers to keep all white space intact as it displays your content (like the following sample). Use the element in place of the element to make the browser apply all your white space, as shown in Figure 5-6.

```
<!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">
<html xmlns="http://www
```

```
with a lot of white space
thrown in for fun (and as a test of course).
</body>
</html>
```



You may want the browser to display white spaces in an HTML page where proper spacing is important, such as

- Code samples
- ✓ Text tables



You can nest $\langle pre \rangle$ elements inside $\langle blockquote \rangle$ elements to carefully control how the lines of quoted text appear on the page.

Line breaks

By default, browsers usually *wrap* text that appears in block elements, such as paragraphs, headings, and block quotes. If a text line reaches the end of a browser window, the next word automatically starts a new line. You can manually control the end of a text line with a *line break* (denoted by the *
br />* element).



76





Function

The $\langle br \rangle$ element is the HTML equivalent of the manual line break that you use in paragraphs and other blocks of text when you're working in a word-processing program. When a browser sees a $\langle br \rangle$, it ends the line there and starts the next line.



The difference between a line break and a paragraph is that a line break doesn't use any special formatting that you can apply at the end or beginning of a paragraph, such as

- Extra vertical space
- First-line indenting

Formatting

The following markup formats the lines of text in a poem with line breaks. The entire poem is described *as a single paragraph*, and the $\langle br \rangle >$ element marks the end of each line:

```
<!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">
 <head>
   <meta http-equiv="Content-Type" content="text/html; charset=ISO-8859-1" />
   <title> Shakespeare in HTML</title>
  </head>
 <body>
  <h1>Shakespeare's Sonnets XVIII: Shall I compare thee to a summer's day? </h1>
   Shall I compare thee to a summer's day? <br />
     Thou art more lovely and more temperate. <br />
     Rough winds do shake the darling buds of May, <br />
     And summer's lease hath all too short a date. <br />
     Sometime too hot the eye of heaven shines, <br />
     And often is his gold complexion dimm'd; <br />
     And every fair from fair sometime declines, <br />
     By chance or nature's changing course untrimm'd; <br />
     But thy eternal summer shall not fade <br />
     Nor lose possession of that fair thou ow'st; <br />
     Nor shall Death brag thou wander'st in his shade, <br />
     When in eternal lines to time thou grow'st: <br />
     So long as men can breathe or eyes can see, <br />
     So long lives this, and this gives life to thee. <br />
   </body>
</html>
```

Figure 5-7 shows how a browser handles each line break. In this example, the poem isn't left-indented because the element replaces the <blockquote> element.



Horizontal rules

The horizontal rule element ($\langle hr \rangle \rangle$) helps you include solid straight lines (*rules*) on your page.



The browser creates the rule based on the hr /> element, so users don't wait for a graphic to download. A horizontal rule is a good option to

- Break your page into logical sections.
- ✓ Separate your headers and footers from the rest of the page.

Formatting

When you include an $\langle hr \rangle \rangle$ element on your page, like the following HTML, the browser replaces it with a line, as shown in Figure 5-8.

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

```
<body>
   This is a paragraph followed by a horizontal rule.
   <hr />
   This is a paragraph preceded by a horizontal rule.
   </body>
</html>
```



A horizontal rule must always sit on a line by itself; you can't add the \leq hr /> element in the middle of a paragraph (or other block element) and expect the rule to just appear in the middle of the block.

Attributes

Four different attributes control the appearance of each horizontal rule:

✓ width: Specifies line width either in *pixels* or by *percentage of display* area width (which we call "the page" in discussion that follows).

For example, a rule can be 50 pixels wide or take 75 percent of the page.

- ✓ size: Specifies the height of the line in pixels. The default is 1 pixel.
- align: Specifies the horizontal alignment of the rule as either left (the default), center, or right.

If you don't define a width for your rule, it takes the entire width of the page. The alignment won't make any difference.

noshade: Specifies a solid line with no shading.

By default, most browsers display hard rules with a shade.



These formatting attributes are deprecated in favor of using CSS.

This bit of HTML creates a horizontal rule that takes up 45 percent of the page, is 4 pixels high, aligned to the center, and has shading turned off:

```
This is a paragraph followed by a horizontal rule.
<hr width="45%" size="4" align="center" noshade="noshade" />
This is a paragraph preceded by a horizontal rule.
```

Figure 5-9 shows how the addition of these attributes can alter how a browser displays the rule.

Figure 5-10 shows how you can use horizontal rules in the real world to highlight important content. The LANWrights, Inc., Web site uses colored hard rules to frame a key statement on the site's home page. The rules make the statement stand out from the rest of the page.











CSS gives you much more control over the placement of horizontal rules; you can even fancy them up with color and shading options.

Organizing Information

Lists are powerful tools for arranging similar elements together, and they give visitors to your site an easy way to hone in on groups of information. You can put just about anything in a list, from a set of instructions to a collection of navigational hyperlinks.

Lists use a combination of elements — at least two components:

- ✓ A markup element that says "Hey browser! The following items are a list."
- Markup elements that say "Hey browser! This is an item in the list."

HTML provides for three different kinds of lists:

- ✓ Numbered lists
- Bulleted lists
- Definition lists

Numbered lists

A *numbered list* consists of at least two items, each prefaced by a number. Usually, a person numbers a list when the order of the items is important.

You use two kinds of elements for a numbered list:

- \checkmark The ordered list element () specifies that this is a numbered list.
- List item elements (<1i>) mark each item in the list.

Formatting

A numbered list with three items requires elements and content in the following order:

- 1.
- 2. <1i>
- 3. Content for the first list item
- 4.
- 5. <1i>
- 6. Content for the second list item
- 7.
- 8. <1i>
- 9. Content for the third list item
- 10.
- 11.

The following markup defines a three-item numbered list:

```
Wash car
Cli>Grocery shopping

</body>
</html>
```

Figure 5-11 shows how a browser renders this markup. You don't actually have to specify a number for each item in the list; the browser identifies the list items from the markup and adds the numbers.

If you swap the first two items in the list, they're still numbered in order when the page appears, as shown in Figure 5-12.

```
Cli>Wash car

Cli>Feed cat

Cli>Grocery shopping
```

Numbering

Two different $\langle 0 \rangle$ element attributes control the appearance of a numbered list:

>> start: Specifies the first number in the list.

• The default starting number is 1.





• You can specify any number as the start number for the new list.

Specify a start number when you resume a list after an unnumbered paragraph or other block element.

type: Specifies the numbering style from the list. You can choose from five predefined numbering styles:

- 1: Decimal numbers.
- a: Lowercase letters.
- A: Uppercase letters.
- i: Lowercase Roman numerals.
- I: Uppercase Roman numerals.

The following markup uses ordered list elements and attributes to create a list that uses uppercase Roman numerals and begins numbering at 5 (V in Roman numerals):

```
   >Wash car
   >Feed cat
   >Grocery shopping
```

Figure 5-13 shows how the attributes affect the list's appearance in a browser.







You have more control over your lists if you use CSS to define formatting. That's why the start and type attributes for list markup are deprecated; see Appendix A for information about deprecated attributes.

Bulleted lists

A *bulleted list* consists of one or more items each prefaced by a *bullet* (often a big dot; this book uses *check marks* as bullets).

You use this type of list if the order of the presentation of the items isn't necessary for understanding the information presented.

Formatting

A bulleted list requires the following:

- The unordered list element () specifies that you're creating a bulleted list.
- ✓ A list item element (<1i>) marks each item in the list.
- The closing tag for the unordered list element () indicates that the list has come to its end.

An unordered list with three items requires elements and content in the following order:

- 1.
- 2. <1i>
- 3. Content for the first list item
- 4.
- 5. <1i>
- 6. Content for the second list item
- 7.
- 8. <1i>

9. Content for the third list item

- 10.
- 11.

The following markup formats a three-item list as a bulleted list:

```
<!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">
 <head>
   <meta http-equiv="Content-Type" content="text/html; charset=ISO-8859-1" />
   <title>Bulleted Lists</title>
 </head>
 <body>
   <h1>Things to do today</h1>
   <u1>
     Feed cat
     Wash car
     Grocery shopping
   </body>
</html>
```

Figure 5-14 shows how a browser renders this with bullets.

Styles

You can use the type attribute (deprecated) with the $\{u\}$ element to specify what kind of bullet you want the list to use.

- disc: Solid circle bullets (the default)
- square: Solid square bullets
- ✓ circle: Hollow circle bullets

	🗿 Bulleted Lists - Microsoft Internet Expl 🔳 🗖
Figure 5-14:	File Edit View Favorites Tools Help 🧸
An	🚱 Back 🔹 🕥 👻 😰 🏠 🔎 Search
unordered	
list uses	Things to do today
bullets	Feed cat
instead of	Freed cat Wash car
numbers to	 Grocery shopping
mark items.	
	Done 🛛 😨 My Computer

The addition of the $\ensuremath{\texttt{type}}$ attribute to the bulleted-list markup just given changes the bullets from discs to squares, as shown in Figure 5-15. Here's what the relevant markup looks like:

W

```
type="square">
Feed cat
Wash car
 Grocery shopping
```





Use CSS if you want more control over the formatting of your lists.

Definition lists

Definition lists group terms and definitions into a single list and require three different elements to complete the list:

- ✓ <d1>: Holds the list definitions.
- dt>: Defines a term in the list.
- dd>: Defines a definition for a term.

You can have as many terms (defined by $\langle dt \rangle$) in a list as you need. Each term can have one or more definitions (defined by $\langle dd \rangle$).

To create a definition list with two items requires elements and content in the following order:

- 1. <dl>
- 2. < dt >
- 3. First term name
- 4.</dt>
- 5. <dd>
- 6. Content for the definition of the first item
- 7. </dd>
- 8.<dt>
- 9. Second term name
- 10.</dt>
- 11. <dd>
- 12. Content for the definition of the second item
- 13. </dd>
- 14.</dl>

The following definition list includes three terms, one of which has two definitions:

```
<title>Definition Lists</title>
</head>
<body>
<h1>Markup Language Definitions</h1>
<dl>
<dt>SGML</dt>
<dd>The Standard Generalized Markup Language</dd>
<dt>HTML</dt>
<dd>The Hypertext Markup Language</dd>
<dd>The Hypertext Markup Language</dd>
<dd>The Hypertext Markup Language</dd>
<dd>The Hypertext Markup Language</dd>
<dd>The Extensible Markup Language</dd>
</dd>
</dd>
```



If you think the items in a list are spaced too closely together, you can either

- Put two
 elements before each or </dd> element to add more white space.
- ✓ Use CSS styles to carefully control all aspects of your list appearance, as shown in Chapter 8.

Nesting lists

You can create subcategories by *nesting* lists within other lists. Some common uses for nested lists include

- ✓ Site maps and other navigation tools
- Table of contents for online books and papers
- 🖌 🖊 Outlines

You can combine any of the three kinds of lists to create *nested* lists, such as a multilevel table of contents or an outline that mixes numbered headings with bulleted list items as the lowest outline level.

The following example starts with a numbered list that defines a list of things to do for the day, and uses three bulleted lists to further break down those items into specific tasks:

```
<!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">
```

```
 <head>
   <meta http-equiv="Content-Type" content="text/html; charset=ISO-8859-1" />
   <title>Nested Lists</title>
 </head>
 <body>
   <h1>Things to do today</h1>
   <01>
    Feed cat
      \langle u \rangle
        Rinse bowl
        Open cat food
        Mix dry and wet food in bowl
        Deliver on a silver platter to fluffy
      Wash car
      \langle u \rangle
        Vacuum interior
        Wash exterior
        Wax exterior
      Grocery shopping
      \langle u \rangle
        Plan meals
        Clean out fridge
        Make list
        Go to store
      </body>
</html>
```

All nested lists follows the same markup pattern:

- Each list item in the top-level ordered list is followed by a complete second-level list.
- ✓ The second-level lists sit inside the top-level list, not in the list items.

Figure 5-16 shows how a browser reflects this nesting in its display.



As you build nested lists, watch your opening and closing tags carefully. *Close first what you opened last* is an especially important axiom here. If you don't open and close your tags properly, lists might not show consistent indents or numbering, or text might be indented incorrectly because a list somewhere was never properly closed.



Text Controls and Annotation

Some general (X)HTML elements define general text controls or allow you to annotate documents. These are covered in Table 5-1.

Table 5-1	(X)HTML Text Controls and Annotation			
Element	Common Name	Empty?	Category	Description
bdo	Bidirectional algorithm	No	Language definition	Controls direction of text display {]tr rt]} (left-to- right, right-to-left)
del	Deleted text	No	Text control	Marks deleted text in current draft
ins	Inserted text	No	Text control	Marks inserted text in current draft

Element	Common Name	Empty?	Category	Description
kbd	Keyboard text	No	Text control	Text to type at a keyboard
samp	Sample text	No	Text control	Sample program output
tt	Teletype text	No	Text control	Typewriter or tele- type output
var	Variable text	No	Text control	Highlights input or output variables

Marvelous Miscellany

Table 5-2 lists other text-related (X)HTML attributes that you might find in HTML files.

Table 5-2	Additional (X)HTML Text Attributes			
Name	Function/ Value Equals	Value Type(s)	Related Element(s)	
cite	Specifies location of source materials	URL	<blockquote> <q></q></blockquote>	
cite	Explains reason for adds, deletes	Text	<ins></ins>	
datetime	Time stamps document content	ISO date	<ins></ins>	
dir	Specifies text direc- tion for content	{ltr rtl}	tl} All elements except <base/> <frame/> <frameset <iframe><param/> <script></td></tr><tr><td>id</td><td>Supplies unique identifier for markup instances</td><td>ID</td><td>All elements except <base /><head> <html><meta /> <param /> <script><style> <title></td></tr></tbody></table></script></iframe></frameset 	

Table 5-2 <i>(continued)</i>				
Name	Function/ Value Equals	Value Type(s)	Related Element(s)	
lang	Names content language used	Language code	All elements except <base/> <frame/> <frameset> <iframe><param/> <script></td></tr><tr><td>title</td><td>Associates advisory info to content</td><td>Text</td><td>Allelementsexcept <base /><head> <html><meta /> <param /><script> <style><title></td></tr></tbody></table></script></br></iframe></frameset>	