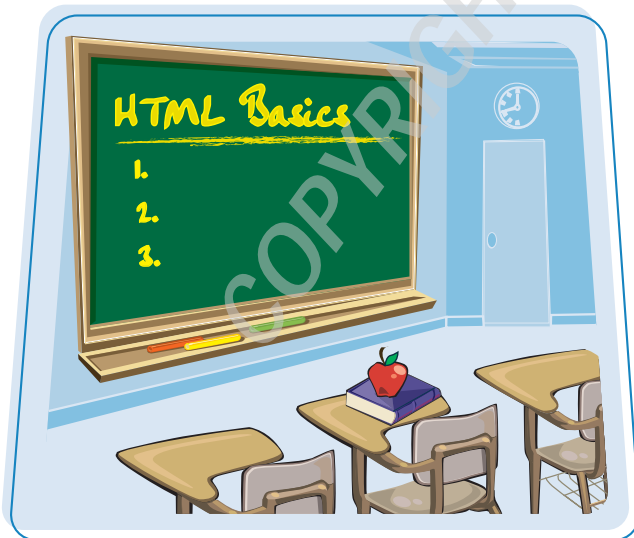


CHAPTER

1

Getting Familiar with HTML and Web Page Basics



Are you interested in building your own Web pages? This chapter introduces you to HTML, the language used to create Web pages. It also explains the basics behind HTML editors and Web browsers, which you use to design and view your Web content.

Internet Basics	4
An Introduction to HTML	6
Explore Web Browsers	8
Explore HTML Editors	9
Understanding HTML Syntax	10
View HTML Code in a Browser.....	12
Plan a Web Site	14

Internet Basics

The Internet is a worldwide collection of interconnected computer networks that enables businesses, organizations, governments, and individuals to communicate in a variety of ways. One of the most popular ways users communicate on the Internet is by publishing and interacting with Web pages. You can also use the Internet to send and receive e-mail, chat with other users, and transfer files between computers.

The Internet began as a military research project in the late 1960s. In 2007, the number of Internet users around the globe topped 1.2 billion.



Types of Connections

Users connect to the Internet through a variety of methods. A relatively inexpensive but slow way to connect is with dialup service, which involves using a modem and a phone line. Faster ways to connect include DSL (digital subscriber line), cable modem, satellite, and ISDN (integrated services digital network). Networks include special wireless transmitters that allow computers to access the Internet wirelessly. Companies that help you connect to the Internet are known as *Internet service providers*, or ISPs.



Connection Speeds

Connection speeds play an important part in a user's Internet experience because slower connections result in slower file transfers and Web page viewing. Dialup connections offer the slowest access to the Internet at up to 56 kilobits per second, or Kbps, followed by ISDN connections at 64 to 128 Kbps. DSL usually offers connection speeds of up to 3 megabits per second, or Mbps, while cable modems can achieve speeds of up to 6 Mbps. A Web page that takes about 20 seconds to download via dialup can take less than a second using a cable modem.



Communication Standards

The Internet infrastructure relies on a variety of protocols that dictate how computers and networks talk to each other. For example, *Transmission Control Protocol/Internet Protocol*, or TCP/IP, is a set of rules that control how Internet messages flow between computers. *HyperText Transfer Protocol*, or HTTP, is a set of rules that determine how browsers should request Web pages and how server computers should deliver them. Having agreed-upon protocols allows seamless communication among the many different types of computers that connect to the Internet.



The World Wide Web

The World Wide Web is a giant collection of documents, or pages, stored on computers around the globe. Commonly called *the Web*, this collection of pages represents a wealth of text, images, audio, and video available to anyone with a computer and an Internet connection. Web pages are stored on servers, which are Internet-connected computers running software that allows them to serve up information to other computers. When you place a text file, image, or other document in a special Web directory on a server, that information is available for other Web users to view. Chapter 15 talks about how to transfer information to a Web server.



URLs and Links

Every page on the Web has a unique address called a URL, which is short for *Uniform Resource Locator*. A URL looks like this:

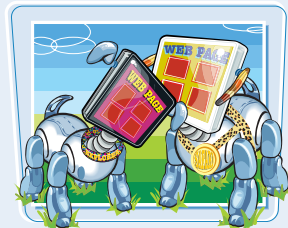
`http://www.example.com/index.html`

If you know a page's URL, you can type it into a Web browser to view that page over the Internet. You can also view pages by way of *hyperlinks*, or simply *links*, which are clickable words or images on Web pages. Every link on a Web page is associated with a URL that leads to another page on the Internet. Users can jump from one Web page to another by clicking links. Chapter 6 discusses how to create links with HTML.



Browsers

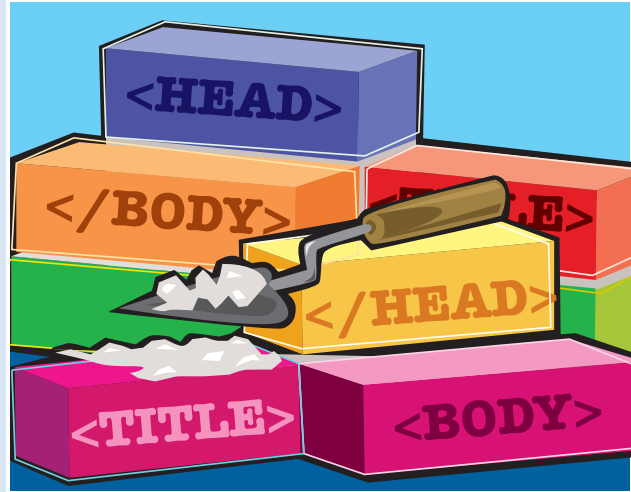
A Web browser is software that allows you to view and interact with Web pages. When you type a URL or click a link in a Web browser, the browser retrieves the appropriate page from a server on the Internet and displays that page. Microsoft Internet Explorer, Mozilla Firefox, and Apple Safari are the three most popular browsers in use today. Each program has evolved through a number of versions, with newer versions supporting more recent Web features. As you build your pages using HTML code, remember that different browsers may display your pages slightly differently depending on the version.



An Introduction to HTML

You build Web pages using HTML, which is short for *HyperText Markup Language*. HTML documents are made up of text content and special codes known as *tags* that tell Web browsers how to display the content. HTML documents are identified by their .html or .htm file extensions.

For the most part, HTML is platform independent, which means you can view Web pages on any computer operating system, including Windows, Mac, and Linux.



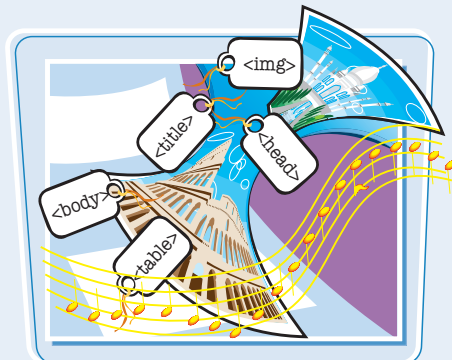
HTML Tags

HTML consists of text interspersed with special instructions known as *tags*. Surrounded by brackets, `<>`, HTML tags tell a browser how to organize and present text, images, and other Web page content. Many tags are written using an opening tag and a closing tag that surround content that appears on the page. When writing HTML tags, you can use upper- or lowercase letters. To make the coding easy to distinguish from other text in the page, you can type tag names in uppercase. For details, see the section “Understanding HTML Syntax.”



Rendering HTML

When a browser displays a Web page, it retrieves the HTML file for that page from a server, parses the HTML tags to determine how the content should be formatted, and renders the page. The HTML tags tell the browser what images, video, audio, and other content need to be downloaded and integrated into the page. The HTML may also tell the browser to download style sheets and interactive scripts to further enhance the page. To view the HTML underlying a Web page, see the section “View HTML Code in a Browser.”



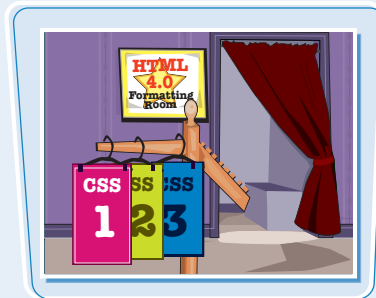
HTML Standards

The World Wide Web Consortium, or W3C, is the primary group guiding the evolution of the HTML language. The W3C is made up of hundreds of companies and organizations including Web industry leaders such as Microsoft, Apple, and Google. The standards developed by the W3C give developers of Web servers and browsers a set of common guidelines with which to develop their products. You can visit the W3C's Web site at www.w3.org.



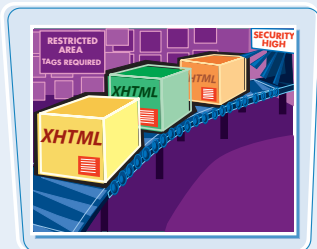
HTML Versions

The most recent version of HTML is 4.01. Version 4.01 includes rules for using more than 90 HTML tags, most of which are covered in this book. It improves on previous versions by adding better support for multimedia, scripting, and style sheets. Support for style sheets is especially important because it allows developers to apply more precise formatting to Web pages. It also allows developers to keep complex styling information separate from the rest of the HTML. Style sheets are covered in Chapters 10, 11, and 12.



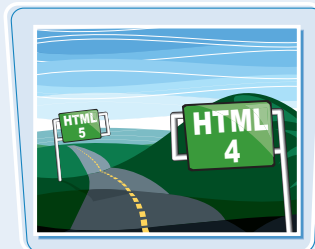
XHTML

XHTML, or *Extensible HyperText Markup Language*, is an alternative language for coding Web pages that conforms to the stricter standards of XML, or *Extensible Markup Language*. XHTML is tag-based and uses many of the same tags as in HTML. However, in XHTML, all tags must be closed, tag names and attributes must be coded in lowercase, and attribute values for tags must be surrounded by quotes. Most modern browsers can read both HTML and XHTML. Although XHTML is not covered in this book, you can read more about it at the W3C site at www.w3.org.



Next Generation of HTML

As this book is being published, the W3C is developing the specification for HTML 5, the next version of HTML. This version will introduce features to help Web designers more easily create and maintain the dynamic pages seen in many of today's Web sites. HTML 5 will include tags for defining headers, footers, and navigation sections, along with tags for adding interactive elements such as editable and sortable tables. For more information, see www.whatwg.org/html5.



Explore Web Browsers

A Web browser is software that can retrieve HTML documents from the Web, parse the HTML instructions, and display the resulting Web pages. You can also use a browser to display HTML documents you save locally on your computer. When coding your HTML, you can use a Web browser to test your work.



Finding a Browser

Most computer operating systems come with a Web browser already installed. Microsoft Windows Vista computers include the Internet Explorer browser, while Apple Mac computers include the Safari browser. (The examples in this book use Internet Explorer.) Mozilla Firefox is another Web browser that has become increasingly popular in recent years. You can learn more about Firefox and download it free at www.mozilla.com/firefox. For more information about the Web browsers in use today, see the Wikipedia at http://wikipedia.org/wiki/Web_browser.



Browser Discrepancies

There are many Web browsers in use today, and numerous versions of each. While most of them interpret HTML essentially the same way, slight differences in interpretation mean that not all of them display Web pages exactly the same way. Also, some more recent browser versions recognize newer HTML features that older browsers do not. You can avoid surprises by writing clean, well-formed HTML code and testing your pages in different browsers as you work. The Wikipedia offers a detailed comparison of Web browser features at http://wikipedia.org/wiki/Comparison_of_web_browsers.

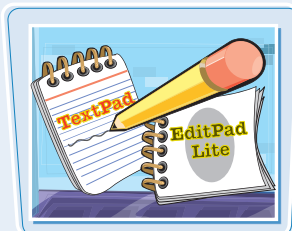


Because HTML documents are plain-text documents, you can use any text-editing program to code HTML and create a Web page. You can also use a variety of Web-specific coding environments that will write your HTML code, validate it, and upload it to a Web server.



Simple Text Editors

Simple text editors, also called plain-text editors, are easy to find. Microsoft Windows Vista comes with Notepad, while Apple Mac computers come with TextEdit. Simple text editors offer no-frills word processing and are often the best choice when you are learning to write HTML. This book uses the Windows WordPad and Notepad text editors in its examples. The Wikipedia has a list of free and commercial text editors at http://wikipedia.org/wiki/List_of_text_editors.



HTML Editors

HTML editors, such as Adobe Dreamweaver and Microsoft Expression, are dedicated programs for writing HTML code and managing Web pages. These programs can shield you from having to write HTML code by offering a graphical environment for building Web pages as well as a text-based environment. Most HTML editors will also color your HTML tags for easier viewing, validate your code, and help you upload finished pages to a server.



Word Processing Programs

You can also use word processing programs, such as Microsoft Word, to write HTML. In Word, you can select HTML as the file type when you save a document, and the program automatically adds the appropriate HTML tags. However, commercial word processors tend to store lots of extra information with your HTML, which can make it a challenge to edit the files in other editors.



Understanding HTML Syntax

HTML is a language for describing Web page content. HTML rules, or *syntax*, govern the way in which code is written. Learning the right way to write your code can save you time and confusion later.



Writing HTML

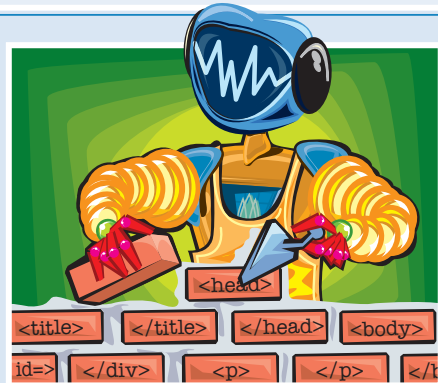
In HTML, tags determine how page content is organized and formatted. Tags consist of words or abbreviations surrounded by angle brackets, `< >`. Tags can be written using upper- or lowercase letters. You can type tag names in uppercase to distinguish the code from other text. This HTML code creates a paragraph in your page:

```
<P>Hello, world!</P>
```



Tag Structure

Certain structural HTML tags identify different parts of your HTML document. For example, the `<BODY>` and `</BODY>` tags surround the main body content that appears in the browser window. Many tags, such as the paragraph tags (`<P>` and `</P>`), are written using an opening tag and a closing tag while others, such as the image tag (``), stand alone. Closing tags must always include a slash (`/`) before the tag name.

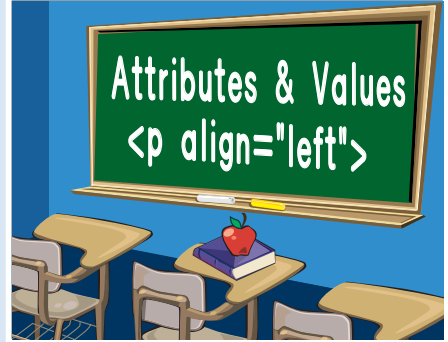


Attributes and Values

You can assign specific attributes to each HTML tag to customize its behavior. Most attributes work by setting a numeric or descriptive value. For example, you can set a paragraph's alignment on the page using the ALIGN attribute along with a type of alignment: left, right, or center. The code for creating a centered paragraph looks like this:

```
<P ALIGN="center">My centered text.</P>
```

Attributes always go inside the opening HTML tag, and it is good form to enclose attribute values in quotation marks.



Entities

You can add special characters to a page, such as a copyright symbol or a fraction, by using special codes called *entities*. Entities represent characters not readily available on the keyboard. All entities are preceded by an ampersand (&) and followed by a semicolon (;). For example, the following code adds a copyright symbol to your page:

```
&copy;
```

For more about entities and special characters, see Chapter 3.

Avoiding Syntax Errors

To avoid HTML errors, always take the time to proofread your code. Most HTML editors have features that highlight bad syntax. Make sure your tags have brackets, your closing tags include a slash, and your attribute values are surrounded by quotation marks. Multiple HTML tags should be properly nested, meaning your closing tags should be in the reverse order of the opening tags. For example:

```
<P ALIGN="center"><B>My text.</B></P>
```

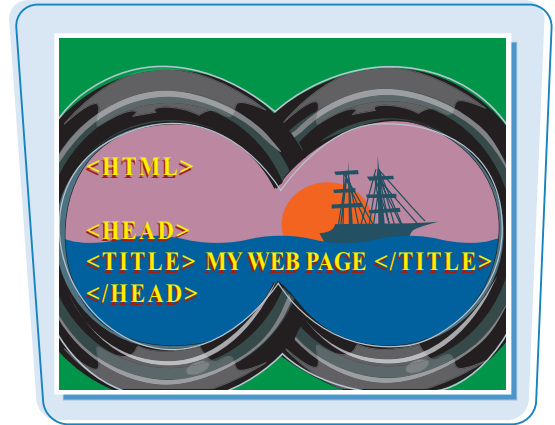
To help make your HTML readable, consider using new lines to type code instead of running everything together on one long line. Doing so will not affect how your page is displayed, because Web browsers ignore extra white space.



View HTML Code in a Browser

You can view the HTML code for any Web page that you have loaded into your Web browser. Viewing HTML from different Web sites is a good way to learn how to write your own code and can spawn new ideas for your own pages. You can also save a Web page locally for use as a template or to study later.

In Microsoft Windows Vista, Internet Explorer opens the HTML code in the Notepad text editor. To view an HTML page that you have saved locally, see Chapter 2.



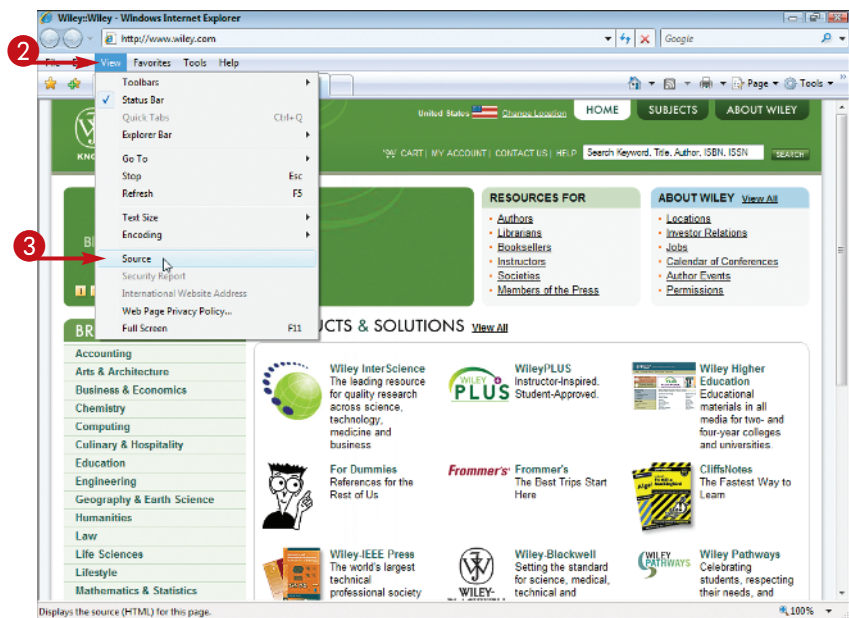
View HTML Code in a Browser

VIEW THE SOURCE CODE


1 Open a Web page in your browser window.

2 Click **View**.

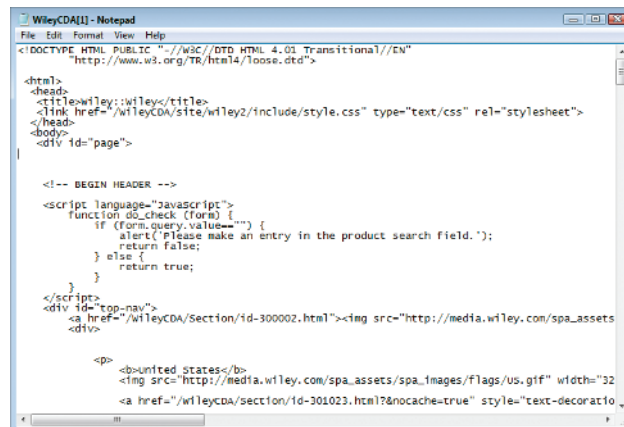
3 Click **Source**.



A Notepad window appears displaying the HTML source code for the page.

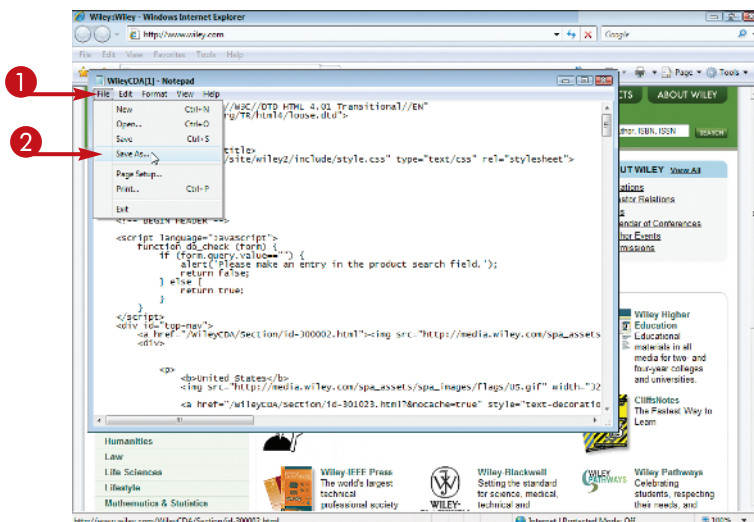
4 Click the Close button  when finished.

The window closes.



SAVE THE SOURCE CODE

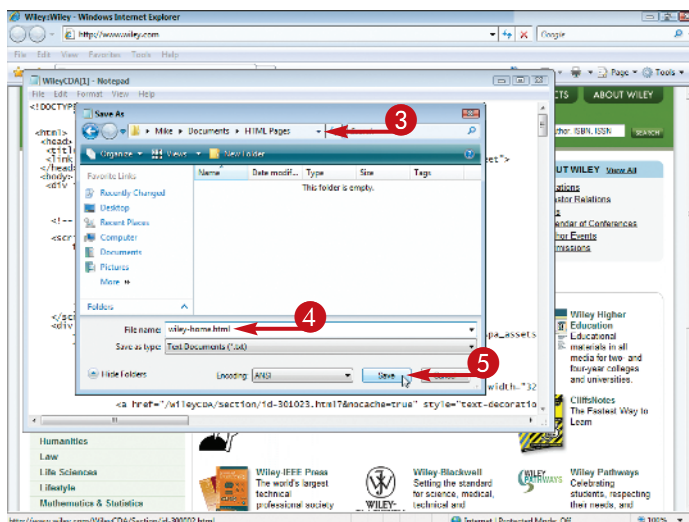
- 1 In the Notepad window that displays the source code, click **File**.
- 2 Click **Save As**.



The Save As dialog box appears.

- 3 Click here to navigate to the folder where you want to store the page.
- 4 Type a name for the page.
- 5 Click **Save**.

HTML pages should have an .html or .htm file extension.



TIPS

Will the HTML documents that I save to my computer work when I open them in a browser?

Possibly. It depends on how the HTML is coded. In addition to the HTML, you may have to download images, style sheets, scripts, and other external content separately, and then edit the HTML so that the page references them correctly. For more about referencing content, see Chapter 6.



How else can I save a Web page in my browser?

In Internet Explorer, you can click **File** and then **Save As**. The browser gives you several ways to save your page. "Complete" saves the HTML and embedded content, "HTML only" saves just the HTML, and "Web Archive" saves the HTML and other content as a single file.



Plan a Web Site

A Web site is a group of related Web pages, all hyperlinked together and hosted on a server. Before you start writing the HTML to create a site's pages, it is a good idea to do some planning.



Know Your Audience

It is important to understand the audience that will visit the pages on your Web site. It is important to know their technical ability so that you can use language they are comfortable with. It is also important to know their interests so that you can present imagery and other content that will keep them interested and at your site. You can conduct interviews with potential users early on to get ideas for your site designs. You can also have users test out your site after it launches to get feedback on how to improve it.



Plan a Home Page

The home page is usually the first page a visitor sees when visiting a Web site, so it is important that it concisely communicate the site's purpose and what information users can expect to find. It should also include prominent links to the other important pages on your site. Optimally, users should be able to see all of this information without having to scroll in their browser window. Clear communication is important on all of your site's pages, but especially on the home page.



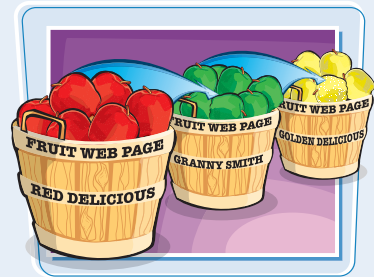
Site Map

A useful tool for planning your site's overall structure is a *site map*, which represents your Web pages as boxes and the hyperlinks as arrows. The home page of a site is usually placed at the top of a site map. A site map gives you an overview of the pages you need to build and also shows the navigational structure. You should arrange your pages so that important content is easy to get to from the home page. You can sketch your site map using pencil and paper or using software such as Microsoft Visio, which has tools specifically made for creating site maps.



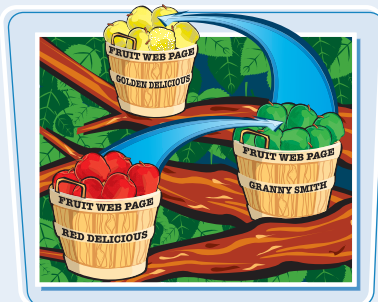
Linear Structure

A linear site layout moves the user through your content in a step-by-step fashion, like pages in a book. Linear layouts are good for presenting sequential instructions or a narrative story. In a linear layout, each Web page usually links to the next page and the previous page. The site map of a linear site will have the pages arranged one after the other in a line.



Hierarchical Structure

A hierarchical layout resembles a pyramid, with the home page at the top and other pages fanning out from there. A hierarchical site map looks like a company's organizational chart or a family tree. Hierarchical layouts are appropriate for sites with categorized content, such as online merchants. Each branch in such a site represents a product category with the for-sale items at the end of the branches.



Gathering Content

After you plan the pages and structure of your site, you need to gather the content. For simple sites, this may involve writing text and shooting digital photos. More complex sites may require recording audio and video, creating illustrations, and programming interactive media. You can organize all of this content into your Web pages using HTML.

