

4.6 Uniform Resource Locators And Hyperlinks

The Web uses a syntactic form known as a *Uniform Resource Locator (URL)* to specify a web page. The general form of a URL is:

```
protocol://computer_name:port/document_name?parameters
```

where *protocol* is the name of the protocol used to access the document, *computer_name* is the domain name of the computer on which the document resides, *:port* is an optional protocol port number at which the server is listening, *document_name* is the optional name of the document on the specified computer, and *%parameters* give optional parameters for the page.

For example, the URL

```
http://www.netbook.cs.purdue.edu/toc/toc01.htm
```

specifies protocol *http*, a computer named *www.netbook.cs.purdue.edu*, and a file named *toc/toc01.htm*.

Typical URLs that a user enters omit many of the parts. For example, the URL:

```
www.netbook.cs.purdue.edu
```

omits the protocol (*http* is assumed), the port (80 is assumed), the document name (*index.html* is assumed), and parameters (none are assumed).

A URL contains the information a browser needs to retrieve a page. The browser uses the separator characters colon, slash, and question mark, to divide the URL into four components: a protocol, a computer name, a document name, and parameters. The browser uses the computer name and protocol port to form a connection to the server on which the page resides, and uses the document name and parameters to request a specific page.

In HTML, an *anchor* tag uses URLs to provide a hyperlink capability (i.e., the ability to link from one web document to another). The following example shows an HTML source document with an anchor surrounding the name *Prentice Hall*:

```
This book is published by  
<A HREF="http://www.prenhall.com">  
Prentice Hall, </A> one of  
the larger publishers of Computer Science textbooks.
```

The anchor references the URL *http://www.prenhall.com*. When displayed on a screen, the HTML input produces:

```
This book is published by Prentice Hall, one of the larger  
publishers of Computer Science textbooks.
```