

PROGRAMMING THE WEB

Subject Code: 10CS73
Hours/Week : 04
Total Hours : 52

I.A. Marks : 25
Exam Hours: 03
Exam Marks: 100

UNIT – 1**6 Hours**

Fundamentals of Web, XHTML – 1: Internet, WWW, Web Browsers and Web Servers, URLs, MIME, HTTP, Security, The Web Programmers Toolbox. XHTML: Basic syntax, Standard structure, Basic text markup, Images, Hypertext Links.

UNIT – 2**7 Hours**

XHTML – 2, CSS: XHTML (continued): Lists, Tables, Forms, Frames

CSS: Introduction, Levels of style sheets, Style specification formats, Selector forms, Property value forms, Font properties, List properties, Color, Alignment of text, The box model, Background images, The and <div> tags, Conflict resolution.

UNIT – 3**6 Hours**

Javascript: Overview of Javascript, Object orientation and Javascript, Syntactic characteristics, Primitives, operations, and expressions, Screen output and keyboard input, Control statements, Object creation and modification, Arrays, Functions, Constructors, Pattern matching using regular expressions, Errors in scripts, Examples.

UNIT – 4**7 Hours**

Javascript and HTML Documents, Dynamic Documents with Javascript: The Javascript execution environment, The Document Object Model, Element access in Javascript, Events and event handling, Handling events from the Body elements, Button elements, Text box and Password elements, The DOM 2 event model, The navigator object, DOM tree traversal and modification.

Introduction to dynamic documents, Positioning elements, Moving elements, Element visibility, Changing colors and fonts, Dynamic content, Stacking elements, Locating the mouse cursor, Reacting to a mouse click, Slow movement of elements, Dragging and dropping elements.

PART - B**UNIT – 5****6 Hours**

XML: Introduction, Syntax, Document structure, Document type definitions, Namespaces, XML schemas, Displaying raw XML documents, Displaying XML documents with CSS, XSLT style sheets, XML processors, Web services.

UNIT – 6**7 Hours**

Perl, CGI Programming: Origins and uses of Perl, Scalars and their operations, Assignment statements and simple input and output, Control statements, Fundamentals of arrays, Hashes, References, Functions, Pattern matching, File input and output; Examples.

The Common Gateway Interface; CGI linkage; Query string format; CGI.pm module; A survey example; Cookies. Database access with Perl and MySQL

UNIT – 7**6 Hours**

PHP: Origins and uses of PHP, Overview of PHP, General syntactic characteristics, Primitives, operations and expressions, Output, Control statements, Arrays, Functions, Pattern matching, Form handling, Files, Cookies, Session tracking, Database access with PHP and MySQL.

UNIT – 8**7 Hours**

Ruby, Rails: Origins and uses of Ruby, Scalar types and their operations, Simple input and output, Control statements, Arrays, Hashes, Methods, Classes, Code blocks and iterators, Pattern matching.

Overview of Rails, Document requests, Processing forms, Rails applications with Databases, Layouts.

Text Books:

1. Robert W. Sebesta: Programming the World Wide Web, 4th Edition, Pearson education, 2008. (Listed topics only from Chapters 1 to 9, 11 to 15)

Reference Books:

1. M. Deitel, P.J. Deitel, A. B. Goldberg: Internet & World Wide Web How to Program, 3rd Edition, Pearson education, 2004.
2. Chris Bates: Web Programming Building Internet Applications, 3rd Edition, Wiley India, 2006.
3. Xue Bai et al: The web Warrior Guide to Web Programming, Thomson, 2003.

INDEX SHEET

UNIT No.	UNIT NAME	PAGE NO.
UNIT 1	Fundamentals of Web, XHTML – 1	4-42
UNIT 2	XHTML – 2	43-58
	CSS	59-83
UNIT 3	JAVASCRIPT	84-105
UNIT 4	JAVASCRIPT AND HTML DOCUMENTS	106-143
	DYNAMIC DOCUMENTS WITH JAVASCRIPT	144-154
UNIT 5	XML	155-180
UNIT 6	PERL, CGI PROGRAMMING	181-237
UNIT-7	PHP	238-252
UNIT-8	Rails and Ruby	253-303

UNIT - 1

Fundamentals of Web, XHTML – 1

1.1 A Brief Introduction to the Internet

1.1.1 Origins

In the 1960s, the U.S. Department of Defense (DoD) became interested in developing a new large-scale computer network. The purposes of this network were communications, program sharing, and remote computer access for researchers working on defense-related contracts. One fundamental requirement was that the network be sufficiently robust so that even if some network nodes were lost to sabotage, war, or some more benign cause, the network would continue to function. The DoD's Advanced Research Projects Agency (ARPA)¹ funded the construction of the first such network, which connected about a dozen ARPA-funded research laboratories and universities. The first node of this network was established at UCLA in 1969. Because it was funded by ARPA, the network was named ARPAnet. Despite the initial intentions, the primary early use of ARPAnet was simple text-based communications through e-mail. Because ARPAnet was available only to laboratories and universities that conducted ARPA-funded research, the great majority of educational institutions were not connected. As a result, a number of other networks were developed during the late 1970s and early 1980s, with BITNET and CSNET among them. BITNET, which is an acronym for Because It's Time Network, began at the City University of New York. It was built initially to provide electronic mail and file transfers. CSNET, which is an acronym for Computer Science Network, connected the University of Delaware, Purdue University, the University of Wisconsin, the RAND Corporation, and Bolt, Beranek, and Newman (a research company in Cambridge, Massachusetts). Its initial purpose was to provide electronic mail. For a variety of reasons, neither BITNET nor CSNET became a widely used national network.

A new national network, NSFnet, was created in 1986. It was sponsored, of course, by the National Science Foundation (NSF). NSFnet initially connected the NSF-funded supercomputer centers at five universities. Soon after being established, it became available to other academic institutions and research laboratories. By 1990, NSFnet had replaced ARPAnet for most nonmilitary uses, and a wide variety of organizations had established nodes on the new network—by 1992 NSFnet connected more than 1 million computers around the world. In 1995, a small part of NSFnet returned to being a research network. The rest became known as the Internet, although this term was used much earlier for both ARPAnet and NSFnet.

1.1.2 What Is the Internet?

The Internet is a huge collection of computers connected in a communications network. These computers are of every imaginable size, configuration, and manufacturer. In fact, some of the devices connected to the Internet—such as plotters and printers—are not computers at all. The innovation that allows all of these diverse devices to communicate with each other is a single, low-level protocol: the Transmission Control Protocol/Internet Protocol (TCP/IP). TCP/IP became the standard for computer network connections in 1982. It can be used directly to allow a program on one computer to communicate with a program on another computer via the Internet. In most cases, however, a higher-level protocol runs on top of TCP/IP. Nevertheless,

it's important to know that TCP/IP provides the low-level interface that allows most computers (and other devices) connected to the Internet to appear exactly the same.²

Rather than connecting every computer on the Internet directly to every other computer on the Internet, normally the individual computers in an organization are connected to each other in a local network. One node on this local network is physically connected to the Internet. So, the Internet is actually a network of networks, rather than a network of computers.

Obviously, all devices connected to the Internet must be uniquely identifiable.

1.1.3 Internet Protocol Addresses

For people, Internet nodes are identified by names; for computers, they are identified by numeric addresses. This relationship exactly parallels the one between a variable name in a program, which is for people, and the variable's numeric memory address, which is for the machine.

The Internet Protocol (IP) address of a machine connected to the Internet is a unique 32-bit number. IP addresses usually are written (and thought of) as four 8-bit numbers, separated by periods. The four parts are separately used by Internet-routing computers to decide where a message must go next to get to its destination.

Organizations are assigned blocks of IPs, which they in turn assign to their machines that need Internet access—which now include most computers. For example, a small organization may be assigned 256 IP addresses, such as 191.57.126.0 to 191.57.126.255. Very large organizations, such as the Department of Defense, may be assigned 16 million IP addresses, which include IP addresses with one particular first 8-bit number, such as 12.0.0.0 to 12.255.255.255.

Although people nearly always type domain names into their browsers, the IP works just as well. For example, the IP for United Airlines (www.ual.com) is 209.87.113.93. So, if a browser is pointed at <http://209.87.113.93>, it will be connected to the United Airlines Web site.

In late 1998, a new IP standard, IPv6, was approved, although it still is not widely used. The most significant change was to expand the address size from 32 bits to 128 bits. This is a change that will soon be essential because the number of remaining unused IP addresses is diminishing rapidly. The new standard can be found at <ftp://ftp.isi.edu/in-notes/rfc2460.txt>.

1.1.4 Domain Names

Because people have difficulty dealing with and remembering numbers, machines on the Internet also have textual names. These names begin with the name of the host machine, followed by progressively larger enclosing collections of machines, called domains. There may be two, three, or more domain names. The first domain name, which appears immediately to the right of the host name, is the domain of which the host is a part. The second domain name gives the domain of which the first domain is a part. The last domain name identifies the type of organization in which the host resides, which is the largest domain in the site's name. For organizations in the United States, edu is the extension for educational institutions, com specifies a company, gov is used for the U.S. government, and org is used for many other kinds of organizations. In other countries, the largest domain is often an abbreviation for the country—for example, se is used for Sweden, and kz is used for Kazakhstan.

Consider this sample address:

movies.comedy.marxbros.com

Here, movies is the hostname and comedy is movies's local domain, which is a part of marxbros's domain, which is a part of the com domain. The hostname and all of the domain names are together called a fully qualified domain name.

Because IP addresses are the addresses used internally by the Internet, the fully qualified domain name of the destination for a message, which is what is given by a browser user, must be converted to an IP address before the message can be transmitted over the Internet to the destination. These conversions are done by software systems called name servers, which implement the Domain Name System (DNS). Name servers serve a collection of machines on the Internet and are operated by organizations that are responsible for the part of the Internet to which those machines are connected. All document requests from browsers are routed to the nearest name server. If the name server can convert the fully qualified domain name to an IP address, it does so. If it cannot, the name server sends the fully qualified domain name to another name server for conversion. Like IP addresses, fully qualified domain names must be unique. Figure 1.1 shows how fully qualified domain names requested by a browser are translated into IPs before they are routed to the appropriate Web server.

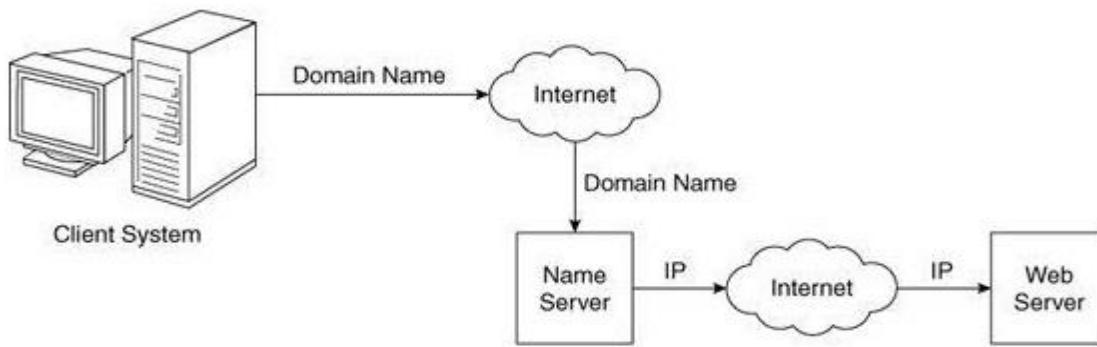


Figure 1.1 Domain name conversion

One way to determine the IP address of a Web site is by using telnet on the fully qualified domain name. This approach is illustrated in Section 1.7.1.

By the mid-1980s, a collection of different protocols that run on top of TCP/IP had been developed to support a variety of Internet uses. Among these protocols, the most common were telnet, which was developed to allow a user on one computer on the Internet to log onto and use another computer on the Internet; File Transfer Protocol (ftp), which was developed to transfer files among computers on the Internet; Usenet, which was developed to serve as an electronic bulletin board; and mailto, which was developed to allow messages to be sent from the user of one computer on the Internet to other users of other computers on the Internet.

This variety of protocols, each with its own user interface and useful only for the purpose for which it was designed, restricted the growth of the Internet. Users were required to learn all the different interfaces to gain all the advantages of the Internet. Before long, however, a better approach was developed: the World Wide Web.

1.2 The World Wide Web

1.2.1 Origins

In 1989, a small group of people led by Tim Berners-Lee at CERN (Conseil Européen pour la Recherche Nucléaire, or European Organization for Particle Physics) proposed a new protocol for the Internet, as well as a system of document access to use it.³ The intent of this new system, which the group named the World Wide Web, was to allow scientists around the world to use the Internet to exchange documents describing their work.

The proposed new system was designed to allow a user anywhere on the Internet to search for and retrieve documents from databases on any number of different document-serving computers connected to the Internet. By late 1990, the basic ideas for the new system had been fully developed and implemented on a NeXT computer at CERN. In 1991, the system was ported to other computer platforms and released to the rest of the world.

For the form of its documents, the system used hypertext, which is text with embedded links to text in other documents to allow nonsequential browsing of textual material. The idea of hypertext had been developed earlier and had appeared in Xerox's NoteCards and Apple's HyperCard in the mid-1980s.

From here on, we will refer to the World Wide Web simply as "the Web." The units of information on the Web have been referred to by several different names; among them, the most common are pages, documents, and resources. Perhaps the best of these is documents, although that seems to imply only text. Pages is widely used, but it is misleading in that Web units of information often have more than one of the kind of pages that make up printed media. There is some merit to calling these units resources, because that covers the possibility of nontextual information. This book will use documents and pages more or less interchangeably, but we prefer documents in most situations.

Documents are sometimes just text, usually with embedded links to other documents, but they often also include images, sound recordings, or other kinds of media. When a document contains nontextual information, it is called hypermedia.

In an abstract sense, the Web is a vast collection of documents, some of which are connected by links. These documents are accessed by Web browsers, introduced in Section 1.3, and are provided by Web servers, introduced in Section 1.4.

1.2.2 Web or Internet?

It is important to understand that the Internet and the Web are not the same thing. The Internet is a collection of computers and other devices connected by equipment that allows them to communicate with each other. The Web is a collection of software and protocols that has been installed on most, if not all, of the computers on the Internet. Some of these computers run Web servers, which provide documents, but most run Web clients, or browsers, which request documents from servers and display them to users. The Internet was quite useful before the Web was developed, and it is still useful without it. However, most users of the Internet now use it through the Web.

1.3 Web Browsers

When two computers communicate over some network, in many cases one acts as a client and

the other as a server. The client initiates the communication, which is often a request for information stored on the server, which then sends that information back to the client. The Web, as well as many other systems, operates in this client-server configuration.

Documents provided by servers on the Web are requested by browsers, which are programs running on client machines. They are called browsers because they allow the user to browse the resources available on servers. The first browsers were text based—they were not capable of displaying graphic information, nor did they have a graphical user interface. This limitation effectively constrained the growth of the Web. In early 1993, things changed with the release of Mosaic, the first browser with a graphical user interface. Mosaic was developed at the National Center for Supercomputer Applications (NCSA) at the University of Illinois. Mosaic's interface provided convenient access to the Web for users who were neither scientists nor software developers. The first release of Mosaic ran on UNIX systems using the X Window system. By late 1993, versions of Mosaic for Apple Macintosh and Microsoft Windows systems had been released. Finally, users of the computers connected to the Internet around the world had a powerful way to access anything on the Web anywhere in the world. The result of this power and convenience was an explosive growth in Web usage.

A browser is a client on the Web because it initiates the communication with a server, which waits for a request from the client before doing anything. In the simplest case, a browser requests a static document from a server. The server locates the document among its servable documents and sends it to the browser, which displays it for the user. However, more complicated situations are common. For example, the server may provide a document that requests input from the user through the browser. After the user supplies the requested input, it is transmitted from the browser to the server, which may use the input to perform some computation and then return a new document to the browser to inform the user of the results of the computation. Sometimes a browser directly requests the execution of a program stored on the server. The output of the program is then returned to the browser.

Although the Web supports a variety of protocols, the most common one is the Hypertext Transfer Protocol (HTTP). HTTP provides a standard form of communication between browsers and Web servers. Section 1.7 presents an introduction to HTTP.

The most commonly used browsers are Microsoft Internet Explorer (IE), which runs only on PCs that use one of the Microsoft Windows operating systems,⁴ and Firefox, which is available in versions for several different computing platforms, including Windows, Mac OS, and Linux. Several other browsers are available, such as the close relatives of Firefox and Netscape Navigator, as well as Opera and Apple's Safari. However, because the great majority of browsers now in use are either IE or Firefox, in this book we focus on those two.

1.4 Web Servers

Web servers are programs that provide documents to requesting browsers. Servers are slave programs: They act only when requests are made to them by browsers running on other computers on the Internet.

The most commonly used Web servers are Apache, which has been implemented for a variety of computer platforms, and Microsoft's Internet Information Server (IIS), which runs under Windows operating systems. As of June 2009, there were over 75 million active Web hosts in operation,⁵ about 47 percent of which were Apache, about 25 percent of which were IIS, and the remainder of which were spread thinly over a large number of others. (The third-place

server was qq.com, a product of a Chinese company, with almost 13 percent.)⁶

1.4.1 Web Server Operation

Although having clients and servers is a natural consequence of information distribution, this configuration offers some additional benefits for the Web. On the one hand, serving information does not take a great deal of time. On the other hand, displaying information on client screens is time consuming. Because Web servers need not be involved in this display process, they can handle many clients. So, it is both a natural and an efficient division of labor to have a small number of servers provide documents to a large number of clients.

Web browsers initiate network communications with servers by sending them URLs (discussed in Section 1.5). A URL can specify one of two different things: the address of a data file stored on the server that is to be sent to the client, or a program stored on the server that the client wants executed, with the output of the program returned to the client.

All the communications between a Web client and a Web server use the standard Web protocol, Hypertext Transfer Protocol (HTTP), which is discussed in Section 1.7.7

When a Web server begins execution, it informs the operating system under which it is running that it is now ready to accept incoming network connections through a specific port on the machine. While in this running state, the server runs as a background process in the operating system environment. A Web client, or browser, opens a network connection to a Web server, sends information requests and possibly data to the server, receives information from the server, and closes the connection. Of course, other machines exist between browsers and servers on the network—specifically, network routers and domain-name servers. This section, however, focuses on just one part of Web communication: the server.

Simply put, the primary task of a Web server is to monitor a communications port on its host machine, accept HTTP commands through that port, and perform the operations specified by the commands. All HTTP commands include a URL, which includes the specification of a host server machine. When the URL is received, it is translated into either a file name (in which case the file is returned to the requesting client) or a program name (in which case the program is run and its output is sent to the requesting client). This process sounds pretty simple, but, as is the case in many other simple-sounding processes, a large number of complicating details are involved.

All current Web servers have a common ancestry: the first two servers, developed at CERN in Europe and NCSA at the University of Illinois. Currently, the most common server configuration is Apache running on some version of UNIX.

1.4.2 General Server Characteristics

Most of the available servers share common characteristics, regardless of their origin or the platform on which they run. This section provides brief descriptions of some of these characteristics.

The file structure of a Web server has two separate directories. The root of one of these is called the document root. The file hierarchy that grows from the document root stores the Web documents to which the server has direct access and normally serves to clients. The root of the other directory is called the server root. This directory, along with its descendant directories,

stores the server and its support software.

The files stored directly in the document root are those available to clients through top-level URLs. Typically, clients do not access the document root directly in URLs; rather, the server maps requested URLs to the document root, whose location is not known to clients. For example, suppose that the site name is www.tunias.com (not a real site—at least, not yet), which we will assume to be a UNIX-based system. Suppose further that the document root is named topdocs and is stored in the /admin/web directory, making its address /admin/web/topdocs. A request for a file from a client with the URL <http://www.tunias.com/petunias.html> will cause the server to search for the file with the file path /admin/web/topdocs/petunias.html. Likewise, the URL <http://www.tunias.com/bulbs/tulips.html> will cause the server to search for the file with the address /admin/web/topdocs/bulbs/tulips.html.

Many servers allow part of the servable document collection to be stored outside the directory at the document root. The secondary areas from which documents can be served are called virtual document trees. For example, the original configuration of a server might have the server store all its servable documents from the primary system disk on the server machine. Later, the collection of servable documents might outgrow that disk, in which case part of the collection could be stored on a secondary disk. This secondary disk might reside on the server machine or on some other machine on a local area network. To support this arrangement, the server is configured to direct-request URLs with a particular file path to a storage area separate from the document-root directory. Sometimes files with different types of content, such as images, are stored outside the document root.

Early servers provided few services other than the basic process of returning requested files or the output of programs whose execution had been requested. The list of additional services has grown steadily over the years. Contemporary servers are large and complex systems that provide a wide variety of client services. Many servers can support more than one site on a computer, potentially reducing the cost of each site and making their maintenance more convenient. Such secondary hosts are called virtual hosts.

Some servers can serve documents that are in the document root of other machines on the Web; in this case, they are called proxy servers.

Although Web servers were originally designed to support only the HTTP protocol, many now support ftp, gopher, news, and mailto. In addition, nearly all Web servers can interact with database systems through Common Gateway Interface (CGI) programs and server-side scripts.

1.4.3 Apache

Apache began as the NCSA server, httpd, with some added features. The name Apache has nothing to do with the Native American tribe of the same name. Rather, it came from the nature of its first version, which was a patchy version of the httpd server. As seen in the usage statistics given at the beginning of this section, Apache is the most widely used Web server. The primary reasons are as follows: Apache is an excellent server because it is both fast and reliable. Furthermore, it is open-source software, which means that it is free and is managed by a large team of volunteers, a process that efficiently and effectively maintains the system. Finally, it is one of the best available servers for Unix-based systems, which are the most popular for Web servers.

Apache is capable of providing a long list of services beyond the basic process of serving

documents to clients. When Apache begins execution, it reads its configuration information from a file and sets its parameters to operate accordingly. A new copy of Apache includes default configuration information for a “typical” operation. The site manager modifies this configuration information to fit his or her particular needs and tastes.

For historical reasons, there are three configuration files in an Apache server: httpd.conf, srm.conf, and access.conf. Only one of these, httpd.conf, actually stores the directives that control an Apache server’s behavior. The other two point to httpd.conf, which is the file that contains the list of directives that specify the server’s operation. These directives are described at <http://httpd.apache.org/docs/2.2/mod/quickreference.html>.

1.4.4 IIS

Although Apache has been ported to the Windows platforms, it is not the most popular server on those systems. Because the Microsoft IIS server is supplied as part of Windows—and because it is a reasonably good server—most Windows-based Web servers use IIS. Apache and IIS provide similar varieties of services.

From the point of view of the site manager, the most important difference between Apache and IIS is that Apache is controlled by a configuration file that is edited by the manager to change Apache’s behavior. With IIS, server behavior is modified by changes made through a window-based management program, named the IIS snap-in, which controls both IIS and ftp. This program allows the site manager to set parameters for the server.

Under Windows XP and Vista, the IIS snap-in is accessed by going to Control Panel, Administrative Tools, and IIS Admin. Clicking on this last selection takes you to a window that allows starting, stopping, or pausing IIS. This same window allows IIS parameters to be changed when the server has been stopped.

1.5 Uniform Resource Locators

Uniform (or universal)8 resource locators (URLs) are used to identify documents (resources) on the Internet. There are many different kinds of resources, identified by different forms of URLs.

1.5.1 URL Formats

All URLs have the same general format:

scheme:object-address

The scheme is often a communications protocol. Common schemes include http, ftp, gopher, telnet, file, mailto, and news. Different schemes use object addresses that have different forms. Our interest here is in the HTTP protocol, which supports the Web. This protocol is used to request and send eXtensible Hypertext Markup Language (XHTML) documents. In the case of HTTP, the form of the object address of a URL is as follows:

//fully-qualified-domain-name/path-to-document

Another scheme of interest to us is file. The file protocol means that the document resides on the machine running the browser. This approach is useful for testing documents to be made

available on the Web without making them visible to any other browser. When file is the protocol, the fully qualified domain name is omitted, making the form of such URLs as follows:

file://path-to-document

Because the focus of this book is on XHTML documents, the remainder of the discussion of URLs is limited to the HTTP protocol.

The host name is the name of the server computer that stores the document (or provides access to it on some other computer). Messages to a host machine must be directed to the appropriate process running on the host for handling. Such processes are identified by their associated port numbers. The default port number of Web server processes is 80. If a server has been configured to use some other port number, it is necessary to attach that port number to the hostname in the URL. For example, if the Web server is configured to use port 800, the host name must have :800 attached.

URLs can never have embedded spaces.⁹ Also, there is a collection of special characters, including semicolons, colons, and ampersands (&), that cannot appear in a URL. To include a space or one of the disallowed special characters, the character must be coded as a percent sign (%) followed by the two-digit hexadecimal ASCII code for the character. For example, if San Jose is a domain name, it must be typed as San%20Jose (20 is the hexadecimal ASCII code for a space). All of the details characterizing URLs can be found at http://www.w3.org/Addressing/URL/URI_Overview.html.

1.5.2 URL Paths

The path to the document for the HTTP protocol is similar to a path to a file or directory in the file system of an operating system and is given by a sequence of directory names and a file name, all separated by whatever separator character the operating system uses. For UNIX servers, the path is specified with forward slashes; for Windows servers, it is specified with backward slashes. Most browsers allow the user to specify the separators incorrectly—for example, using forward slashes in a path to a document file on a Windows server, as in the following:

http://www.gumboco.com/files/f99/storefront.html

The path in a URL can differ from a path to a file because a URL need not include all directories on the path. A path that includes all directories along the way is called a complete path. In most cases, the path to the document is relative to some base path that is specified in the configuration files of the server. Such paths are called partial paths. For example, if the server's configuration specifies that the root directory for files it can serve is files/f99, the previous URL is specified as follows:

http://www.gumboco.com/storefront.html

If the specified document is a directory rather than a single document, the directory's name is followed immediately by a slash, as in the following:

<http://www.gumboco.com/departments/>

Sometimes a directory is specified (with the trailing slash) but its name is not given, as in the following example:

<http://www.gumboco.com/>

The server then searches at the top level of the directory in which servable documents are normally stored for something it recognizes as a home page. By convention, this page is often a file named index.html. The home page usually includes links that allow the user to find the other related servable files on the server.

If the directory does not have a file that the server recognizes as being a home page, a directory listing is constructed and returned to the browser.

1.6 Multipurpose Internet Mail Extensions

A browser needs some way of determining the format of a document it receives from a Web server. Without knowing the form of the document, the browser would be unable to render it, because different document formats require different rendering tools. The forms of these documents are specified with Multipurpose Internet Mail Extensions (MIME).

1.6.1 Type Specifications

MIME was developed to specify the format of different kinds of documents to be sent via Internet mail. These documents could contain various kinds of text, video data, or sound data. Because the Web has needs similar to those of Internet mail, MIME was adopted as the way to specify document types transmitted over the Web. A Web server attaches a MIME format specification to the beginning of the document that it is about to provide to a browser. When the browser receives the document from a Web server, it uses the included MIME format specification to determine what to do with the document. If the content is text, for example, the MIME code tells the browser that it is text and also indicates the particular kind of text it is. If the content is sound, the MIME code tells the browser that it is sound and then gives the particular representation of sound so that the browser can choose a program to which it has access to produce the transmitted sound.

MIME specifications have the following form:

type/subtype

The most common MIME types are text, image, and video. The most common text subtypes are plain and html. Some common image subtypes are gif and jpeg. Some common video subtypes are mpeg and quicktime. A list of MIME specifications is stored in the configuration files of every Web server. In the remainder of this book, when we say document type, we mean both the document's type and its subtype.

Servers determine the type of a document by using the filename's extension as the key into a table of types. For example, the extension .html tells the server that it should attach text/html to the document before sending it to the requesting browser.¹⁰

Browsers also maintain a conversion table for looking up the type of a document by its file name extension. However, this table is used only when the server does not specify a MIME

type, which may be the case with some older servers. In all other cases, the browser gets the document type from the MIME header provided by the server.

1.6.2 Experimental Document Types

Experimental subtypes are sometimes used. The name of an experimental subtype begins with x-, as in video/x-msvideo. Any Web provider can add an experimental subtype by having its name added to the list of MIME specifications stored in the Web provider's server. For example, a Web provider might have a handcrafted database whose contents he or she wants to make available to others through the Web. Of course, this raises the issue of how the browser can display the database. As might be expected, the Web provider must supply a program that the browser can call when it needs to display the contents of the database. These programs either are external to the browser, in which case they are called helper applications, or are code modules that are inserted into the browser, in which case they are called plug-ins.

Every browser has a set of MIME specifications (file types) it can handle. All can deal with text/plain (unformatted text) and text/html (HTML files), among others. Sometimes a particular browser cannot handle a specific document type, even though the type is widely used. These cases are handled in the same way as the experimental types described previously. The browser determines the helper application or plug-in it needs by examining the browser configuration file, which provides an association between file types and their required helpers or plug-ins. If the browser does not have an application or a plug-in that it needs to render a document, an error message is displayed.

1.7 The Hypertext Transfer Protocol

All Web communications transactions use the same protocol: the Hypertext Transfer Protocol (HTTP). The current version of HTTP is 1.1, formally defined as RFC 2616, which was approved in June 1999. RFC 2616 is available at the Web site for the World Wide Web Consortium (W3C), <http://www.w3.org>. This section provides a brief introduction to HTTP.

HTTP consists of two phases: the request and the response. Each HTTP communication (request or response) between a browser and a Web server consists of two parts: a header and a body. The header contains information about the communication; the body contains the data of the communication if there is any.

1.7.1 The Request Phase

The general form of an HTTP request is as follows:

1. HTTP method Domain part of the URL HTTP version
2. Header fields
3. Blank line
4. Message body

The following is an example of the first line of an HTTP request:

GET /storefront.html HTTP/1.1

Only a few request methods are defined by HTTP, and even a smaller number of these are typically used. Table 1.1 lists the most commonly used methods.

Table 1.1 HTTP request methods

Method	Description
GET	Returns the contents of the specified document
HEAD	Returns the header information for the specified document
POST	Executes the specified document, using the enclosed data
PUT	Replaces the specified document with the enclosed data
DELETE	Deletes the specified document

Among the methods given in Table 1.1, GET and POST are the most frequently used. POST was originally designed for tasks such as posting a news article to a newsgroup. Its most common use now is to send form data from a browser to a server, along with a request to execute a program on the server that will process the data.

Following the first line of an HTTP communication is any number of header fields, most of which are optional. The format of a header field is the field name followed by a colon and the value of the field. There are four categories of header fields:

1. General: For general information, such as the date
2. Request: Included in request headers
3. Response: For response headers
4. Entity: Used in both request and response headers

One common request field is the Accept field, which specifies a preference of the browser for the MIME type of the requested document. More than one Accept field can be specified if the browser is willing to accept documents in more than one format. For example; we might have any of the following:

Accept: text/plain
 Accept: text/html
 Accept: image/gif

A wildcard character, the asterisk (*), can be used to specify that part of a MIME type can be anything. For example, if any kind of text is acceptable, the Accept field could be as follows:

Accept: text/*

The Host: host name request field gives the name of the host. The Host field is required for HTTP 1.1. The If-Modified-Since: date request field specifies that the requested file should be

sent only if it has been modified since the given date.

If the request has a body, the length of that body must be given with a Content-length field, which gives the length of the response body in bytes. POST method requests require this field because they send data to the server.

The header of a request must be followed by a blank line, which is used to separate the header from the body of the request. Requests that use the GET, HEAD, and DELETE methods do not have bodies. In these cases, the blank line signals the end of the request.

A browser is not necessary to communicate with a Web server; telnet can be used instead. Consider the following command, given at the command line of any widely used operating system:

```
> telnet blanca.uccs.edu http
```

This command creates a connection to the http port on the blanca.uccs.edu server. The server responds with the following:¹¹

```
Trying 128.198.162.60 ...
Connected to blanca
Escape character is '^]'.
```

The connection to the server is now complete, and HTTP commands such as the following can be given:

```
GET /~user1/respond.html HTTP/1.1
Host: blanca.uccs.edu
```

1.7.2 The Response Phase

The general form of an HTTP response is as follows:

1. Status line

2. Response header fields

3. Blank line

4. Response body

The status line includes the HTTP version used, a three-digit status code for the response, and a short textual explanation of the status code. For example, most responses begin with the following:

HTTP/1.1 200 OK

The status codes begin with 1, 2, 3, 4, or 5. The general meanings of the five categories specified by these first digits are shown in Table 1.2.

Table 1.2 First digits of HTTP status codes

First Digit	Category
1	Informational
2	Success
3	Redirection
4	Client error
5	Server error

One of the more common status codes is one users never want to see: 404 Not Found, which means the requested file could not be found. Of course, 200 OK is what users want to see, because it means that the request was handled without error. The 500 code means that the server has encountered a problem and was not able to fulfill the request.

After the status line, the server sends a response header, which can contain several lines of information about the response, each in the form of a field. The only essential field of the header is Content-type.

The following is the response header for the request given near the end of Section 1.7.1:

```
HTTP/1.1 200 OK
Date: Sat, 25 July 2009 22:15:11 GMT
Server: Apache/2.2.3 (CentOS)
Last-modified: Tues, 18 May 2004 16:38:38 GMT
ETag: "1b48098-16c-3dab592dc9f80"
Accept-ranges: bytes
Content-length: 364
Connection: close
Content-type: text/html, charset=UTF-8
```

The response header must be followed by a blank line, as is the case for request headers. The response data follows the blank line. In the preceding example, the response body would be the HTML file, respond.html.

In HTTP versions prior to 1.1, when a server finished sending a response to the client, the communications connection was closed. However, the default operation of HTTP 1.1 is that the connection is kept open for a time so that the client can make several requests over a short span of time without needing to reestablish the communications connection with the server. This change led to significant increases in the efficiency of the Web.

1.8 Security

It does not take a great deal of contemplation to realize that the Internet and the Web are fertile grounds for security problems. On the Web server side, anyone on the planet with a computer, a browser, and an Internet connection can request the execution of software on any server computer. He or she can also access data and databases stored on the server computer. On the browser end, the problem is similar: Any server to which the browser points can download software that is to be executed on the browser host machine. Such software can access parts of the memory and memory devices attached to that machine that are not related to the needs of the original browser request. In effect, on both ends, it is like allowing any number of total

strangers into your house and trying to prevent them from leaving anything in the house, taking anything from the house, or altering anything in the house. The larger and more complex the design of the house, the more difficult it will be to prevent any of those activities. The same is true for Web servers and browsers: The more complex they are, the more difficult it is to prevent security breaches. Today's browsers and Web servers are indeed large and complex software systems, so security is a significant problem in Web applications.

The subject of Internet and Web security is extensive and complicated, so much so that more than a few books that focus on it have been written. Therefore, this one section of one chapter of one book can give no more than a brief sketch of some of the subtopics of security.

One aspect of Web security is the matter of getting one's data from the browser to the server and having the server deliver data back to the browser without anyone or any device intercepting or corrupting those data along the way. Consider just the simplest case, that of transmitting a credit card number to a company from which a purchase is being made. The security issues for this transaction are as follows:

1. Privacy—it must not be possible for the credit card number to be stolen on its way to the company's server.
2. Integrity—it must not be possible for the credit card number to be modified on its way to the company's server.
3. Authentication—it must be possible for both the purchaser and the seller to be certain of each other's identity.
4. Nonrepudiation—it must be possible to prove legally that the message was actually sent and received.

The basic tool to support privacy and integrity is encryption. Data to be transmitted is converted into a different form, or encrypted, such that someone (or some computer) who is not supposed to access the data cannot decrypt it. So, if data is intercepted while en route between Internet nodes, the interceptor cannot use the data because he or she cannot decrypt it. Both encryption and decryption are done with a key and a process (applying the key to the data). Encryption was developed long before the Internet ever existed. Julius Caesar crudely encrypted the messages he sent to his field generals while at war. Until the middle 1970s, the same key was used for both encryption and decryption, so the initial problem was how to transmit the key from the sender to the receiver.

This problem was solved in 1976 by Whitfield Diffie and Martin Hellman of Stanford University, who developed public-key encryption, a process in which a public key and a private key are used, respectively, to encrypt and decrypt messages. A communicator—say, Joe—has an inversely related pair of keys, one public and one private. The public key can be distributed to all organizations that might send Joe messages. All of them can use the public key to encrypt messages to Joe, who can decrypt the messages with his matching private key. This arrangement works because the private key need never be transmitted and also because it is virtually impossible to decrypt the private key from its corresponding public key. The technical wording for this situation is that it is “computationally infeasible” to determine the private key from its public key.

The most widely used public-key algorithm is named RSA, developed in 1977 by three MIT professors—Ron Rivest, Adi Shamir, and Leonard Adleman—the first letters of whose last names were used to name the algorithm. Most large companies now use RSA for e-commerce. Another, completely different security problem for the Web is the intentional and malicious destruction of data on computers attached to the Internet. The number of different ways this can be done has increased steadily over the life span of the Web. The sheer number of such attacks has also grown rapidly. There is now a continuous stream of new and increasingly devious denial-of-service (DoS) attacks, viruses, and worms being discovered, which have caused billions of dollars of damage, primarily to businesses that use the Web heavily. Of course, huge damage also has been done to home computer systems through Web intrusions.

DoS attacks can be created simply by flooding a Web server with requests, overwhelming its ability to operate effectively. Most DoS attacks are conducted with the use of networks of virally infected “zombie” computers, whose owners are unaware of their sinister use. So, DoS and viruses are often related.

Viruses are programs that often arrive in a system in attachments to e-mail messages or attached to free downloaded programs. Then they attach to other programs. When executed, they replicate and can overwrite memory and attached memory devices, destroying programs and data alike. Two viruses that were extensively destructive appeared in 2000 and 2001: the ILOVEYOU virus and the CodeRed virus, respectively.

Worms damage memory, like viruses, but spread on their own, rather than being attached to other files. Perhaps the most famous worm so far has been the Blaster worm, spawned in 2003. DoS, virus, and worm attacks are created by malicious people referred to as hackers. The incentive for these people apparently is simply the feeling of pride and accomplishment they derive from being able to cause huge amounts of damage by outwitting the designers of Web software systems.

Protection against viruses and worms is provided by antivirus software, which must be updated frequently so that it can detect and protect against the continuous stream of new viruses and worms.

1.9 The Web Programmer’s Toolbox

This section provides an overview of the most common tools used in Web programming—some are programming languages, some are not. The tools discussed are XHTML, a markup language, along with a few high-level markup document-editing systems; XML, a meta-markup language; JavaScript, PHP, and Ruby, which are programming languages; JSF, ASP.NET, and Rails, which are development frameworks for Web-based systems; Flash, a technology for creating and displaying graphics and animation in XHTML documents; and Ajax, a Web technology that uses JavaScript and XML.

Web programs and scripts are divided into two categories—client side and server side—according to where they are interpreted or executed. XHTML and XML are client-side languages; PHP and Ruby are server-side languages; JavaScript is most often a client-side language, although it can be used for both.

We begin with the most basic tool: XHTML.

1.9.1 Overview of XHTML

At the onset, it is important to realize that XHTML is not a programming language—it cannot be used to describe computations. Its purpose is to describe the general form and layout of documents to be displayed by a browser.

The word markup comes from the publishing world, where it is used to describe what production people do with a manuscript to specify to a printer how the text, graphics, and other elements in the book should appear in printed form. XHTML is not the first markup language used with computers. TeX and LaTeX are older markup languages for use with text; they are now used primarily to specify how mathematical expressions and formulas should appear in print.

An XHTML document is a mixture of content and controls. The controls are specified by the tags of XHTML. The name of a tag specifies the category of its content. Most XHTML tags consist of a pair of syntactic markers that are used to delimit particular kinds of content. The pair of tags and their content together are called an element. For example, a paragraph element specifies that its content, which appears between its opening tag, `<p>`, and its closing tag, `</p>`, is a paragraph. A browser has a default style (font, font style, font size, and so forth) for paragraphs, which is used to display the content of a paragraph element.

Some tags include attribute specifications that provide some additional information for the browser. In the following example, the `src` attribute specifies the location of the `img` tag's image content:

```
<img src = "redhead.jpg"/>
```

In this case, the image document stored in the file `redhead.jpg` is to be displayed at the position in the document in which the tag appears.

XHTML 1.0 was introduced in early 2000 by the W3C as an alternative to HTML 4.01, which was at that time (and still is) the latest version of HTML. XHTML 1.0 is nothing more than HTML 4.01 with stronger syntactic rules. These stronger rules are those of XML (see Section 1.9.4). The current version, XHTML 1.1, was released in May 2001 as a replacement for XHTML 1.0, although, for various reasons, XHTML 1.0 is still widely used. Chapter 2, “Introduction to XHTML,” provides a description of a large subset of XHTML.

1.9.2 Tools for Creating XHTML Documents

XHTML documents can be created with a general-purpose text editor. There are two kinds of tools that can simplify this task: XHTML editors and what-you-see-is-what-you-get (WYSIWYG, pronounced wizzy-wig) XHTML editors.

XHTML editors provide shortcuts for producing repetitious tags such as those used to create the rows of a table. They also may provide a spell-checker and a syntax-checker, and they may color code the XHTML in the display to make it easier to read and edit.

A more powerful tool for creating XHTML documents is a WYSIWYG XHTML editor. Using a WYSIWYG XHTML editor, the writer can see the formatted document that the XHTML describes while he or she is writing the XHTML code. WYSIWYG XHTML editors are very useful for beginners who want to create simple documents without learning XHTML and for users who want to prototype the appearance of a document. Still, these editors sometimes

produce poor-quality XHTML. In some cases, they create proprietary tags that some browsers will not recognize.

Two examples of WYSIWYG XHTML editors are Microsoft FrontPage and Adobe Dreamweaver. Both allow the user to create XHTML-described documents without requiring the user to know XHTML. They cannot handle all of the tags of XHTML, but they are very useful for creating many of the common features of documents. Between the two, FrontPage is by far the most widely used. Information on Dreamweaver is available at <http://www.adobe.com/>; information on FrontPage is available at <http://www.microsoft.com/frontpage/>.

1.9.3 Plug-ins and Filters

Two different kinds of converters can be used to create XHTML documents. Plug-ins¹² are programs that can be integrated together with a word processor. Plug-ins add new capabilities to the word processor, such as toolbar buttons and menu elements that provide convenient ways to insert XHTML into the document being created or edited. The plug-in makes the word processor appear to be an XHTML editor that provides WYSIWYG XHTML document development. The end result of this process is an XHTML document. The plug-in also makes available all the tools that are inherent in the word processor during XHTML document creation, such as a spell-checker and a thesaurus.

A second kind of converter is a filter, which converts an existing document in some form, such as LaTeX or Microsoft Word, to XHTML. Filters are never part of the editor or word processor that created the document—an advantage because the filter can then be platform independent. For example, a Word-Perfect user working on a Macintosh computer can use a filter running on a UNIX platform to provide documents that can be later converted to XHTML. The disadvantage of filters is that creating XHTML documents with a filter is a two-step process: First you create the document, and then you use a filter to convert it to XHTML.

Neither plugs-ins nor filters produce XHTML documents that, when displayed by browsers, have the identical appearance of that produced by the word processor.

The two advantages of both plug-ins and filters, however, are that existing documents produced with word processors can be easily converted to XHTML and that users can use a word processor with which they are familiar to produce XHTML documents. This obviates the need to learn to format text by using XHTML directly. For example, once you learn to create tables with your word processor, it is easier to use that process than to learn to define tables directly in XHTML.

The XHTML output produced by either filters or plug-ins often must be modified, usually with a simple text editor, to perfect the appearance of the displayed document in the browser. Because this new XHTML file cannot be converted to its original form (regardless of how it was created), you will have two different source files for a document, inevitably leading to version problems during maintenance of the document. This is clearly a disadvantage of using converters.

1.9.4 Overview of XML

HTML is defined with the use of the Standard Generalized Markup Language (SGML), which is a language for defining markup languages. (Such languages are called meta-markup

languages.) XML (eXtensible Markup Language) is a simplified version of SGML, designed to allow users to easily create markup languages that fit their own needs. XHTML is defined with the use of XML. Whereas XHTML users must use the predefined set of tags and attributes, when a user creates his or her own markup language with XML, the set of tags and attributes is designed for the application at hand. For example, if a group of users wants a markup language to describe data about weather phenomena, that language could have tags for cloud forms, thunderstorms, and low-pressure centers. The content of these tags would be restricted to relevant data. If such data is described with XHTML, cloud forms could be put in generic tags, but then they could not be distinguished from thunderstorm elements, which would also be in the same generic tags.

Whereas XHTML describes the overall layout and gives some presentation hints for general information, XML-based markup languages describe data and its meaning through their individualized tags and attributes. XML does not specify any presentation details.

The great advantage of XML is that application programs can be written to use the meanings of the tags in the given markup language to find specific kinds of data and process it accordingly. The syntax rules of XML, along with the syntax rules for a specific XML-based markup language, allow documents to be validated before any application attempts to process their data. This means that all documents that use a specific markup language can be checked to determine whether they are in the standard form for such documents. Such an approach greatly simplifies the development of application programs that process the data in XML documents.

1.9.5 Overview of JavaScript

JavaScript is a client-side scripting language whose primary uses in Web programming are to validate form data and to create dynamic XHTML documents.

The name JavaScript is misleading because the relationship between Java and JavaScript is tenuous, except for some of the syntax. One of the most important differences between JavaScript and most common programming languages is that JavaScript is dynamically typed. This strategy is virtually the opposite of that of strongly typed languages such as C++ and Java. JavaScript “programs” are usually embedded in XHTML documents,¹³ which are downloaded from a Web server when they are requested by browsers. The JavaScript code in an XHTML document is interpreted by an interpreter embedded in the browser on the client.

One of the most important applications of JavaScript is to dynamically create and modify documents. JavaScript defines an object hierarchy that matches a hierarchical model of an XHTML document. Elements of an XHTML document are accessed through these objects, providing the basis for dynamic documents.

Chapter 4, “The Basics of JavaScript,” provides a more detailed look at JavaScript. Chapter 5, “JavaScript and XHTML Documents,” and Chapter 6, “Dynamic Documents with JavaScript,” discuss the use of JavaScript to provide access to, and dynamic modification of, XHTML documents.

1.9.6 Overview of Flash

There are two components of Flash: the authoring environment, which is a development framework, and the player. Developers use the authoring environment to create static graphics, animated graphics, text, sound, and interactivity to be part of stand-alone HTML documents or

to be part of other XHTML documents. These documents are served by Web servers to browsers, which use the Flash player plug-in to display the documents. Much of this development is done by clicking buttons, choosing menu items, and dragging and dropping graphics.

Flash makes animation very easy. For example, for motion animation, the developer needs only to supply the beginning and ending positions of the figure to be animated—Flash builds the intervening figures. The interactivity of a Flash application is implemented with ActionScript, a dialect of JavaScript.

Flash is now the leading technology for delivering graphics and animation on the Web. It has been estimated that nearly 99 percent of the world's computers used to access the Internet have a version of the Flash player installed as a plug-in in their browsers.

1.9.7 Overview of PHP

PHP is a server-side scripting language specifically designed for Web applications. PHP code is embedded in XHTML documents, as is the case with JavaScript. With PHP, however, the code is interpreted on the server before the XHTML document is delivered to the requesting client. A requested document that includes PHP code is preprocessed to interpret the PHP code and insert its output into the XHTML document. The browser never sees the embedded PHP code and is not aware that a requested document originally included such code.

PHP is similar to JavaScript, both in terms of its syntactic appearance and in terms of the dynamic nature of its strings and arrays. Both JavaScript and PHP use dynamic data typing, meaning that the type of a variable is controlled by the most recent assignment to it. PHP's arrays are a combination of dynamic arrays and hashes (associative arrays). The language includes a large number of predefined functions for manipulating arrays.

PHP allows simple access to XHTML form data, so form processing is easy with PHP. PHP also provides support for many different database management systems. This versatility makes it an excellent language for building programs that need Web access to databases.

1.9.8 Overview of Ajax

Ajax, shorthand for Asynchronous JavaScript + XML, has been around for a few years, but did not acquire its catchy name until 2005.¹⁴ The idea of Ajax is relatively simple, but it results in a different way of viewing and building Web interactions. This new approach produces an enriched Web experience for those using a certain category of Web interactions.

In a traditional (as opposed to Ajax) Web interaction, the user sends messages to the server either by clicking a link or by clicking a form's Submit button. After the link has been clicked or the form has been submitted, the client waits until the server responds with a new document. The entire browser display is then replaced by the new document. Complicated documents take a significant amount of time to be transmitted from the server to the client and more time to be rendered by the browser. In Web applications that require frequent interactions with the client and remain active for a significant amount of time, the delay in receiving and rendering a complete response document can be disruptive to the user.

In an Ajax Web application, there are two variations from the traditional Web interaction. First, the communication from the browser to the server is asynchronous; that is, the browser need not wait for the server to respond. Instead, the browser user can continue whatever he or she was doing while the server finds and transmits the requested document and the browser renders

the new document. Second, the document provided by the server usually is only a relatively small part of the displayed document, and therefore it takes less time to be transmitted and rendered. These two changes can result in much faster interactions between the browser and the server.

The x in Ajax, from XML, is there because in many cases the data supplied by the server is in the form of an XML document, which provides the new data to be placed in the displayed document. However, in some cases the data is plain text, which may even be JavaScript code. It can also be XHTML.

The goal of Ajax is to have Web-based applications become closer to desktop (client-resident) applications, in terms of the speed of interactions and the quality of the user experience. Wouldn't we all like our Web-based applications to be as responsive as our word processors?

Ajax has some advantages over the competing technologies of ASP.NET and JSP. First and foremost, the technologies that support Ajax are already resident in nearly all Web browsers and servers. This is in contrast to both ASP.NET and JSP, which still have far-from-complete coverage. Second, using Ajax does not require learning a new tool or language. Rather, it requires only a new way of thinking about Web interactions.

Ajax is discussed in more depth in Chapter 10, "Introduction to Ajax."

1.9.9 Overview of Servlets, JavaServer Pages, and JavaServer Faces

There are many computational tasks in a Web interaction that must occur on the server, such as processing order forms and accessing server-resident databases. A Java class called a servlet can be used for these applications. A servlet is a compiled Java class, an object of which is executed on the server system when requested by the XHTML document being displayed by the browser. A servlet produces an XHTML document as a response, some parts of which are static and are generated by simple output statements, while other parts are created dynamically when the servlet is called.

When an HTTP request is received by a Web server, the Web server examines the request. If a servlet must be called, the Web server passes the request to the servlet processor, called a servlet container. The servlet container determines which servlet must be executed, makes sure that it is loaded, and calls it. As the servlet handles the request, it generates an XHTML document as its response, which is returned to the server through the response object parameter.

Java can also be used as a server-side scripting language. An XHTML document with embedded Java scriptlets is one form of JavaServer Pages (JSP). Built on top of servlets, JSP provides alternative ways of constructing dynamic Web documents. JSP takes an opposite approach to that of servlets: Instead of embedding XHTML in Java code that provides dynamic documents, code of some form is embedded in XHTML documents to provide the dynamic parts of a document. These different forms of code make up the different approaches used by JSP. The basic capabilities of servlets and JSP are the same.

When requested by a browser, a JSP document is processed by a software system called a JSP container. Some JSP containers compile the document when the document is loaded on the server; others compile it only when requested. The compilation process translates a JSP document into a servlet and then compiles the servlet. So, JSP is actually a simplified approach to writing servlets.

JavaServer Faces (JSF) adds another layer to the JSP technology. The most important

contribution of JSF is an event-driven user interface model for Web applications. Client-generated events can be handled by server-side code with JSF.

Servlets, JSP, and JSF are discussed in Chapter 11, “Java Web Software.”

1.9.10 Overview of Active Server Pages .NET

Active Server Pages .NET (ASP.NET) is a Microsoft framework for building server-side dynamic documents. ASP.NET documents are supported by programming code executed on the Web server. As we saw in Section 1.9.9, JSF uses Java to describe the dynamic generation of XHTML documents, as well as to describe computations associated with user interactions with documents. ASP.NET provides an alternative to JSF, with two major differences: First, ASP.NET allows the server-side programming code to be written in any of the .NET languages.¹⁵ Second, in ASP.NET all programming code is compiled, which allows it to execute much faster than interpreted code.

Every ASP.NET document is compiled into a class. From a programmer’s point of view, developing dynamic Web documents (and the supporting code) in ASP.NET is similar to developing non-Web applications. Both involve defining classes based on library classes, implementing interfaces from a library, and calling methods defined in library classes. An application class uses, and interacts with, existing classes. In ASP.NET, this is exactly the same for Web applications: Web documents are designed by designing classes.

ASP.NET is discussed in Chapter 12, “Introduction to ASP.NET.”

1.9.11 Overview of Ruby

Ruby is an object-oriented interpreted scripting language designed by Yukihiro Matsumoto (a.k.a. Matz) in the early 1990s and released in 1996. Since then, it has continually evolved and its level of usage has grown rapidly. The original motivation for Ruby was dissatisfaction of its designer with the earlier languages Perl and Python.

The primary characterizing feature of Ruby is that it is a pure object-oriented language, just as is Smalltalk. Every data value is an object and all operations are via method calls. The operators in Ruby are only syntactic mechanisms to specify method calls for the corresponding operations. Because they are methods, many of the operators can be redefined by user programs. All classes, whether predefined or user defined, can have subclasses.

Both classes and objects in Ruby are dynamic in the sense that methods can be dynamically added to either. This means that classes and objects can have different sets of methods at different times during execution. So, different instantiations of the same class can behave differently.

The syntax of Ruby is related to that of Eiffel and Ada. There is no need to declare variables, because dynamic typing is used. In fact, all variables are references and do not have types, although the objects they reference do.

Our interest in Ruby is based on Ruby’s use with the Web development framework Rails (see Section 1.9.12). Rails was designed for use with Ruby, and it is Ruby’s primary use in Web programming. Programming in Ruby is introduced in Chapter 14, “Introduction to Ruby.”

Ruby is culturally interesting because it is the first programming language designed in Japan that has achieved relatively widespread use outside that country.

1.9.12 Overview of Rails

Rails is a development framework for Web-based applications that access databases. A framework is a system in which much of the more-or-less standard software parts are furnished by the framework, so they need not be written by the applications developer. ASP.NET and JSF are also development frameworks for Web-based applications. Rails, whose more official name is Ruby on Rails, was developed by David Heinemeier Hansson in the early 2000s and was released to the public in July 2004. Since then, it has rapidly gained widespread interest and usage. Rails is based on the Model–View–Controller (MVC) architecture for applications, which clearly separates applications into three parts: presentation, data model, and program logic.

Rails applications are tightly bound to relational databases. Many Web applications are closely integrated with database access, so the Rails–relational-database framework is a widely applicable architecture.

Rails can be, and often is, used in conjunction with Ajax. Rails uses the Java-Script framework Prototype to support Ajax and interactions with the JavaScript model of the document being displayed by the browser. Rails also provides other support for developing Ajax, including producing visual effects.

Rails was designed to be used with Ruby and makes use of the strengths of that language. Furthermore, Rails is written in Ruby. Using Rails is introduced in Chapter 15, “Introduction to Rails.”

1.9 XHTML: Origins and evolution of HTML and XHTML

HTML

What is HTML?

HTML is a language for describing web pages.

- HTML stands for **Hyper Text Markup Language**
- HTML is not a programming language, it is a **markup language**
- A markup language is a set of **markup tags**
- HTML uses **markup tags** to describe web pages

HTML Tags

HTML markup tags are usually called HTML tags

- HTML tags are keywords surrounded by **angle brackets** like <html>
- HTML tags normally **come in pairs** like and
- The first tag in a pair is the **start tag**, the second tag is the **end tag**
- Start and end tags are also called **opening tags** and **closing tags**

Origins

Tim Berners-Lee

In 1980, physicist Tim Berners-Lee, who was a contractor at CERN, proposed and prototyped ENQUIRE, a system for CERN researchers to use and share documents. In 1989, Berners-Lee wrote a memo proposing an Internet-based hypertext system.^[2] Berners-Lee specified HTML and wrote the browser and server software in the last part of 1990. In that year, Berners-Lee and CERN data systems engineer Robert Cailliau collaborated on a joint request for funding, but the project was not formally adopted by CERN. In his personal notes,^[3] from 1990 he lists^[4] "some of the many areas in which hypertext is used", and puts an encyclopedia first.

First specifications

The first publicly available description of HTML was a document called HTML Tags, first mentioned on the Internet by Berners-Lee in late 1991.^{[5][6]} It describes 20 elements comprising the initial, relatively simple design of HTML. Except for the hyperlink tag, these were strongly influenced by SGMLguid, an in-house SGML based documentation format at CERN. Thirteen of these elements still exist in HTML 4.^[7]

HTML is a text and image formatting language used by web browsers to dynamically format web pages. Many of the text elements are found in the 1988 ISO technical report TR 9537 Techniques for using SGML, which in turn covers the features of early text formatting languages such as that used by the RUNOFF command developed in the early 1960s for the CTSS (Compatible Time-Sharing System) operating system: these formatting commands were derived from the commands used by typesetters to manually format documents. However the SGML concept of generalized markup is based on elements (nested annotated ranges with attributes) rather than merely point effects, and also the separation of structure and processing: HTML has been progressively moved in this direction with CSS.

Berners-Lee considered HTML to be an application of SGML, and it was formally defined as such by the Internet Engineering Task Force (IETF) with the mid-1993 publication of the first proposal for an HTML specification: "Hypertext Markup Language (HTML)" Internet-Draft by Berners-Lee and Dan Connolly, which included an SGML Document Type Definition to define the grammar.^[8] The draft expired after six months, but was notable for its acknowledgment of the NCSA Mosaic browser's custom

tag for embedding in-line images, reflecting the IETF's philosophy of basing standards on successful prototypes.^[9] Similarly, Dave Raggett's competing Internet-Draft, "HTML+ (Hypertext Markup Format)", from late 1993, suggested standardizing already-implemented features like tables and fill-out forms.^[10]

After the HTML and HTML+ drafts expired in early 1994, the IETF created an HTML Working Group, which in 1995 completed "HTML 2.0", the first HTML specification intended to be treated as a standard against which future implementations should be based.^[9] Published as Request for Comments 1866, HTML 2.0 included ideas from the HTML and HTML+ drafts.^[11] The 2.0 designation was intended to distinguish the new edition from previous drafts.^[12]

Further development under the auspices of the IETF was stalled by competing interests. Since 1996, the HTML specifications have been maintained, with input from commercial software vendors, by the World Wide Web Consortium (W3C).^[13] However, in 2000, HTML also became an international standard (ISO/IEC 15445:2000). The last HTML specification published by the W3C is the HTML 4.01 Recommendation, published in late 1999. Its issues and errors were last acknowledged by errata published in 2001.

Version history of the standard

HTML version timeline

November 24, 1995

HTML 2.0 was published as IETF RFC 1866. Supplemental RFCs added capabilities:

- November 25, 1995: RFC 1867 (form-based file upload)
- May 1996: RFC 1942 (tables)
- August 1996: RFC 1980 (client-side image maps)
- January 1997: RFC 2070 (internationalization)

In June 2000, all of these were declared obsolete/historic by RFC 2854.

January 1997

HTML 3.2^[14] was published as a W3C Recommendation. It was the first version developed and standardized exclusively by the W3C, as the IETF had closed its HTML Working Group in September 1996.^[15]

HTML 3.2 dropped math formulas entirely, reconciled overlap among various proprietary extensions, and adopted most of Netscape's visual markup tags.

Netscape's blink element and Microsoft's marquee element were omitted due to a mutual agreement between the two companies.^[13] A markup for mathematical formulas similar to that in HTML wasn't standardized until 14 months later in MathML.

December 1997

HTML 4.0^[16] was published as a W3C Recommendation. It offers three variations:

- Strict, in which deprecated elements are forbidden,
- Transitional, in which deprecated elements are allowed,
- Frameset, in which mostly only frame related elements are allowed;

Initially code-named "Cougar",^[17] HTML 4.0 adopted many browser-specific element types and attributes, but at the same time sought to phase out Netscape's visual markup features by marking them as deprecated in favor of style sheets.

HTML 4 is an SGML application conforming to ISO 8879 - SGML.^[18]

April 1998

HTML 4.0^[19] was reissued with minor edits without incrementing the version number.

December 1999

HTML 4.01^[20] was published as a W3C Recommendation. It offers the same three variations as HTML 4.0, and its last errata were published May 12, 2001.

May 2000

ISO/IEC 15445:2000^{[21][22]} ("ISO HTML", based on HTML 4.01 Strict) was published as an ISO/IEC international standard. In the ISO this standard falls in the domain of the ISO/IEC JTC1/SC34 (ISO/IEC Joint Technical Committee 1, Subcommittee 34 - Document description and processing languages).^[21]

As of mid-2008, HTML 4.01 and ISO/IEC 15445:2000 are the most recent versions of HTML. Development of the parallel, XML-based language XHTML occupied the W3C's HTML Working Group through the early and mid-2000s.

HTML draft version timeline

October 1991

HTML Tags,^[5] an informal CERN document listing twelve HTML tags, was first mentioned in public.

June 1992

First informal draft of the HTML DTD, with seven^[citation needed] subsequent revisions

November 1992

HTML DTD 1.1 (the first with a version number, based on RCS revisions, which start with 1.1 rather than 1.0), an informal draft

June 1993

Hypertext Markup Language was published by the IETF IIIR Working Group as an Internet-Draft (a rough proposal for a standard). It was replaced by a second version [1] one month later, followed by six further drafts published by IETF itself [2] that finally led to HTML 2.0 in RFC1866

November 1993

HTML+ was published by the IETF as an Internet-Draft and was a competing proposal to the Hypertext Markup Language draft. It expired in May 1994.

April 1995 (authored March 1995)

HTML 3.0 was proposed as a standard to the IETF, but the proposal expired five months later without further action. It included many of the capabilities that were in Raggett's HTML+ proposal, such as support for tables, text flow around figures, and the display of complex mathematical formulas.^[25]

W3C began development of its own Arena browser for testing support for HTML 3 and Cascading Style Sheets, but HTML 3.0 did not succeed for several reasons. The draft was considered very large at 150 pages and the pace of browser development, as well as the number of interested parties, had outstripped the resources of the IETF.^[13] Browser vendors, including Microsoft and Netscape at the time, chose to implement different subsets of HTML 3's draft features as well as to introduce their own extensions to it.^[13] (See Browser wars) These included extensions to control stylistic aspects of documents, contrary to the "belief [of the academic engineering community] that such things as text color, background texture, font size and font face were definitely outside the scope of a language when their only intent was to specify how a document would be organized." Dave Raggett, who has been a W3C Fellow for many years has commented for example, "To a certain extent, Microsoft built its business on the Web by extending HTML features."

January 2008

HTML 5 was published as a Working Draft by the W3C.

Although its syntax closely resembles that of SGML, HTML 5 has abandoned any attempt to be an SGML application, and has explicitly defined its own "html" serialization, in addition to an alternative XML-based XHTML 5 serialization.^[27]

XHTML versions

Main article: XHTML

XHTML is a separate language that began as a reformulation of HTML 4.01 using XML

1.0. It continues to be developed:

- XHTML 1.0, published January 26, 2000 as a W3C Recommendation, later revised and republished August 1, 2002. It offers the same three variations as HTML 4.0 and 4.01, reformulated in XML, with minor restrictions.
- XHTML 1.1, published May 31, 2001 as a W3C Recommendation. It is based on XHTML 1.0 Strict, but includes minor changes, can be customized, and is reformulated using modules from Modularization of XHTML, which was published April 10, 2001 as a W3C Recommendation.
- XHTML 2.0,. There is no XHTML 2.0 standard. XHTML 2.0 is incompatible with XHTML 1.x and, therefore, would be more accurate to characterize as an XHTML-inspired new language than an update to XHTML 1.x.
- XHTML 5, which is an update to XHTML 1.x, is being defined alongside HTML 5 in the HTML 5 draft.

XHTML

XHTML (Extensible Hypertext Markup Language) is a family of XML markup languages that mirror or extend versions of the widely used Hypertext Markup Language (HTML), the language in which web pages are written.

While HTML (prior to HTML5) was defined as an application of Standard Generalized Markup Language (SGML), a very flexible markup language framework, XHTML is an application of XML, a more restrictive subset of SGML. Because XHTML documents need to be well-formed, they can be parsed using standard XML parsers—unlike HTML, which requires a lenient HTML-specific parser.

XHTML 1.0 became a World Wide Web Consortium (W3C) Recommendation on

January 26, 2000. XHTML 1.1 became a W3C Recommendation on May 31, 2001. XHTML5 is undergoing development as of September 2009, as part of the HTML5 specification.

Origin

XHTML 1.0 is "a reformulation of the three HTML 4 document types as applications of XML 1.0". The World Wide Web Consortium (W3C) also continues to maintain the HTML 4.01 Recommendation, and the specifications for HTML5 and XHTML5 are being actively developed. In the current XHTML 1.0 Recommendation document, as published and revised to August 2002, the W3C commented that, "The XHTML family is the next step in the evolution of the Internet. By migrating to XHTML today, content developers can enter the XML world with all of its attendant benefits, while still remaining confident in their content's backward and future compatibility."^[1]

However, in 2004, the Web Hypertext Application Technology Working Group (WHATWG) formed, independently of the W3C, to work on advancing ordinary HTML not based on XHTML. Most major browser vendors were unwilling to implement the features in new W3C XHTML 2.0 drafts, and felt that they didn't serve the needs of modern web development. The WHATWG eventually began working on a standard that supported both XML and non-XML serializations, HTML 5, in parallel to W3C standards such as XHTML 2. In 2007, the W3C's HTML working group voted to officially recognize HTML 5 and work on it as the next-generated HTML standard. In 2009, the W3C allowed the XHTML 2 Working Group's charter to expire, acknowledging that HTML 5 would be the sole next-generation HTML standard, including both XML and non-XML serializations.

Evolution

XHTML 1.0

December 1998 saw the publication of a W3C Working Draft entitled Reformulating HTML in XML. This introduced Voyager, the codename for a new markup language based on HTML 4 but adhering to the stricter syntax rules of XML. By February 1999 the specification had changed name to XHTML 1.0: The Extensible HyperText Markup Language, and in January 2000 it was officially adopted as a W3C Recommendation.^[29]

There are three formal DTDs for XHTML 1.0, corresponding to the three different versions of HTML 4.01:

- **XHTML 1.0 Strict** is the XML equivalent to strict HTML 4.01, and includes elements and attributes that have not been marked deprecated in the HTML 4.01 specification.
- **XHTML 1.0 Transitional** is the XML equivalent of HTML 4.01 Transitional, and includes the presentational elements (such as center, font and strike) excluded from the strict version.
- **XHTML 1.0 Frameset** is the XML equivalent of HTML 4.01 Frameset, and allows for the definition of frameset documents—a common Web feature in the late 1990s.

The second edition of XHTML 1.0 became a W3C Recommendation in August 2002.^[30]

Modularization of XHTML

Modularization provides an abstract collection of components through which XHTML can be subsetted and extended. The feature is intended to help XHTML extend its reach onto emerging platforms, such as mobile devices and Web-enabled televisions. The initial draft of Modularization of XHTML became available in April 1999, and reached Recommendation status in April 2001.

The first XHTML Family Markup Languages to be developed with this technique were XHTML 1.1 and XHTML Basic 1.0. Another example is XHTML-Print (W3C Recommendation, September 2006), a language designed for printing from mobile devices to low-cost printers.

In October 2008 Modularization of XHTML was superseded by XHTML Modularization 1.1, which adds an XML Schema implementation

XHTML 1.1—Module-based XHTML

XHTML 1.1 evolved out of the work surrounding the initial Modularization of XHTML specification. The W3C released a first draft in September 1999; Recommendation status was reached in May 2001. The modules combined within XHTML 1.1 effectively recreate XHTML 1.0 Strict, with the addition of ruby annotation elements (ruby, rbc, rtc, rb, rt and rp) to better support East-Asian languages. Other changes include removal of the lang attribute (in favour of xml:lang), and removal of the name attribute from the a

and map elements.

Although XHTML 1.1 is largely compatible with XHTML 1.0 and HTML 4, in August 2002 the HTML WG (renamed to XHTML2 WG since) issued a Working Group Note advising that it should not be transmitted with the HTML media type. With limited browser support for the alternate application/xhtml+xml media type, XHTML 1.1 proved unable to gain widespread use. In January 2009 a second edition of the document (XHTML Media Types - Second Edition, not to be confused with the XHTML 1.1 - 2nd ed) was issued, relaxing this restriction and allowing XHTML 1.1 to be served as text/html.

XHTML 1.1 Second Edition (W3C Proposed Edited Recommendation) was issued on 7 May 2009 and rescinded on 19 May 2009. (This does not affect the text/html media type usage for XHTML 1.1 as specified in the: XHTML Media Types - Second Edition)

XHTML Basic and XHTML-MP

To support constrained devices, XHTML Basic was created by the W3C; it reached Recommendation status in December 2000. XHTML Basic 1.0 is the most restrictive version of XHTML, providing a minimal set of features that even the most limited devices can be expected to support.

The Open Mobile Alliance and its predecessor the WAP Forum released three specifications between 2001 and 2006 that extended XHTML Basic 1.0. Known as XHTML Mobile Profile or XHTML-MP, they were strongly focused on uniting the differing markup languages used on mobile handsets at the time. All provide richer form controls than XHTML Basic 1.0, along with varying levels of scripting support.

XHTML Basic 1.1 became a W3C Recommendation in July 2008, superseding XHTML-MP 1.2. XHTML Basic 1.1 is almost but not quite a subset of regular XHTML 1.1. The most notable addition over XHTML 1.1 is the inputmode attribute—also found in XHTML-MP 1.2—which provides hints to help browsers improve form entry.

XHTML 1.2

The XHTML 2 Working Group is considering the creation of a new language based on XHTML 1.1. If XHTML 1.2 is created, it will include WAI-ARIA and role attributes to better support accessible web applications, and improved Semantic Web support through RDFa. The inputmode attribute from XHTML Basic 1.1, along with the target attribute (for specifying frame targets) may also be present. It's important to note that the

XHTML2 WG have not yet been chartered to carry out the development of XHTML1.2 and the W3C has announced that it does not intend to recharter the XHTML2 WG, this means that the XHTML1.2 proposal may not eventuate.

XHTML 2.0

Between August 2002 and July 2006 the W3C released the first eight Working Drafts of XHTML 2.0, a new version of XHTML able to make a clean break from the past by discarding the requirement of backward compatibility. This lack of compatibility with XHTML 1.x and HTML 4 caused some early controversy in the web developer community. Some parts of the language (such as the role and RDFa attributes) were subsequently split out of the specification and worked on as separate modules, partially to help make the transition from XHTML 1.x to XHTML 2.0 smoother. A ninth draft of XHTML 2.0 was expected to appear in 2009, but on July 2, 2009, the W3C decided to let the XHTML2 Working Group charter expire by that year's end, effectively halting any further development of the draft into a standard.

New features introduced by XHTML 2.0 include:

- HTML forms will be replaced by XForms, an XML-based user input specification allowing forms to be displayed appropriately for different rendering devices.
- HTML frames will be replaced by XFrames.
- The DOM Events will be replaced by XML Events, which uses the XML Document Object Model.
- A new list element type, the nl element type, will be included to specifically designate a list as a navigation list. This will be useful in creating nested menus, which are currently created by a wide variety of means like nested unordered lists or nested definition lists.
- Any element will be able to act as a hyperlink, e.g., `<li href="articles.html">Articles`, similar to XLink. However, XLink itself is not compatible with XHTML due to design differences.
- Any element will be able to reference alternative media with the src attribute, e.g., `<p src="lbridge.jpg" type="image/jpeg">London Bridge</p>` is the same as `<object src="lbridge.jpg" type="image/jpeg"><p>London Bridge</p></object>`.
- The alt attribute of the img element has been removed: alternative text will be given in the content of the img element, much like the object element, e.g., `<img`

src="hms_audacious.jpg">HMS Audacious.

- A single heading element (h) will be added. The level of these headings are determined by the depth of the nesting. This allows the use of headings to be infinite, rather than limiting use to six levels deep.
- The remaining presentational elements i, b and tt, still allowed in XHTML 1.x (even Strict), will be absent from XHTML 2.0. The only somewhat presentational elements remaining will be sup and sub for superscript and subscript respectively, because they have significant non-presentational uses and are required by certain languages. All other tags are meant to be semantic instead (e. g. for **strong or bolded** text) while allowing the user agent to control the presentation of elements via CSS.
- The addition of RDF triple with the property and about attributes to facilitate the conversion from XHTML to RDF/XML.

HTML5—Vocabulary and APIs for HTML5 and XHTML5

Main article: HTML5

HTML5 initially grew independently of the W3C, through a loose group of browser manufacturers and other interested parties calling themselves the WHATWG, or Web Hypertext Application Technology Working Group. The WHATWG announced the existence of an open mailing list in June 2004, along with a website bearing the strapline —Maintaining and evolving HTML since 2004.|| The key motive of the group was to create a platform for dynamic web applications; they considered XHTML 2.0 to be too document-centric, and not suitable for the creation of internet forum sites or online shops.

In April 2007, the Mozilla Foundation and Opera Software joined Apple in requesting that the newly rechartered HTML Working Group of the W3C adopt the work, under the name of HTML 5. The group resolved to do this the following month, and the First Public Working Draft of HTML 5 was issued by the W3C in January 2008. The most recent W3C Working Draft was published in June 2008.

HTML5 has both a regular text/html serialization and an XML serialization, which is known as XHTML5. In addition to the markup language, the specification includes a number of application programming interfaces. The Document Object Model is extended

with APIs for editing, drag-and-drop, data storage and network communication.

The language is more compatible with HTML 4 and XHTML 1.x than XHTML 2.0, due to the decision to keep the existing HTML form elements and events model. It adds many new elements not found in XHTML 1.x, however, such as section and aside. (The XHTML 1.2 equivalent (which (X)HTML5 replaces) of these structural elements would be <div role="region"> and <div role="complementary">.)

As of 2009-09-03, the latest editor's draft includes WAI-ARIA support.

1.11 Basic syntax

Writing valid HTML (or XHTML) is not a terribly difficult task once you know what the rules are, although the rules are slightly more stringent in XHTML than in HTML. The list below provides a quick reference to the rules that will ensure your markup is well-formed and valid. Note that there are other differences between HTML and XHTML which go beyond simple syntax requirements; those differences are covered in HTML Versus XHTML.

The Document Tree

A web page is, at its heart, little more than a collection of HTML elements—the defining structures that signify a paragraph, a table, a table cell, a quote, and so on. The element is created by writing an opening tag, and completed by writing a closing tag. In the case of a paragraph, you'd create a p element by typing <p>Content goes here</p>.

The elements in a web page are contained in a tree structure in which html is the root element that splits into the head and body elements (as explained in Basic Structure of a Web Page). An element may contain other nested elements (although this very much depends on what the parent element is; for example, a p element can contain span, em, or strong elements, among others). Where this occurs, the opening and closing tags must be symmetrical. If an opening paragraph tag is followed by the opening em element, the closing tags must appear in the reverse order, like so: <p>Content goes here, and some of it needs emphasis too</p>. If you were to type <p>Content goes here, and some of it needs emphasis too</p>, you'd have created invalid markup.

Case Sensitivity

In HTML, tag names are case insensitive, but in XHTML they're case sensitive. As such, in HTML, you can write the markup in lowercase, mixed case, or uppercase letters. So <p>this is a paragraph</p>, as is <P>this example</P>, and even <P>this markup would be valid</p>. In XHTML, however, you must use lowercase for markup: <p>This is a valid paragraph in XHTML</p>.

Opening and Closing Tags

In HTML, it's possible to omit some closing tags (check each element's reference to see whether an HTML closing tag is required), so this is valid markup: <p>This is my first paragraph.<p>This is my second paragraph.<p>And here's the last one..

In XHTML, all elements must be closed. Hence the paragraph example above would need to be changed to: <p>This is my first paragraph.</p><p>This is my second paragraph.</p><p>And here's the last one.</p>

As well as letting you omit some closing tags, HTML allows you to omit start tags—but only on the html, head, body, and tbody elements. This is not a recommended practice, but is technically possible.

For empty elements such as img, XHTML (that is not served with the application/xhtml+xml) requires us to use the XML empty element syntax: <elementname attribute="attributevalue"/>

If serving the document as application/xhtml+xml, it's also valid to close empty elements using a start and end tag, for example the img element, as

Readability Considerations

A browser doesn't care whether you use a single space to separate attributes, ten spaces, or even complete line breaks; it doesn't matter, as long as some space is present. As such, all of the examples below are perfectly acceptable (although the more spaces you include, the larger your web page's file size will be—each occurrence of whitespace takes up additional bytes—so the first example is still the most preferable):

```
  
  

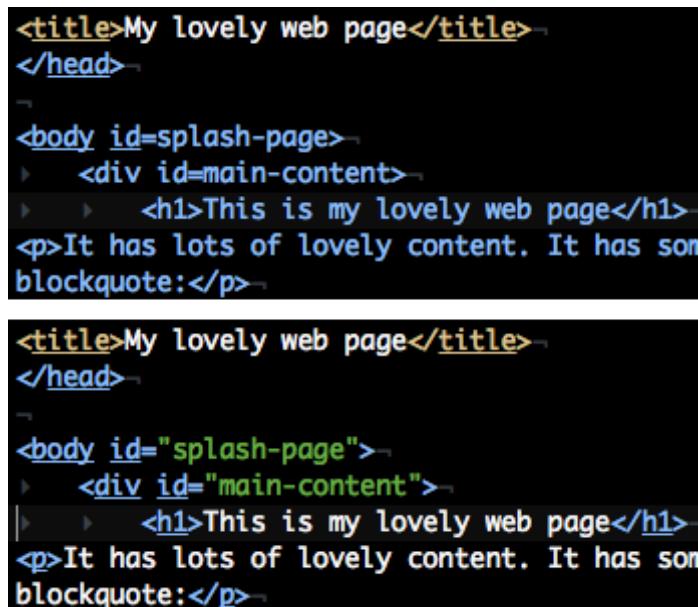

```

In XHTML all attribute values must be quoted, so you'll need to write class="gallery" rather than class=gallery. It's valid to omit the quotes from your HTML, though it may make reading the markup more difficult for developers revisiting old markup (although this really depends on the developer—it's a subjective thing). It's simply easier always to add quotes, rather than to have to remember in which scenarios attribute values require quotes in HTML, as the following piece of HTML demonstrates:

```
<a href="http://example.org"> needs to be quoted because it contains a /
<a href=index.html> acceptable without quotes in HTML
```

Another reason why it's a good idea always to quote your attributes, even if you're using HTML 4.01, is that your HTML editor may be able to provide syntax coloring that makes the code even easier to scan through. Without the quotes, the software may not be able to identify the difference between elements, attributes, and attribute values. This fact is illustrated in Figure , which shows a comparison between quoted and unquoted syntax coloring in the Mac text editor TextMate.

Figure. TextMate's syntax coloring taking effect to display quoted attributes



The figure consists of two side-by-side screenshots of the TextMate text editor. Both screenshots show the same HTML code with different syntax coloring applied based on whether attributes are quoted or unquoted.

Top Screenshot (Quoted Attributes):

```
<title>My lovely web page</title>
</head>

<body id=splash-page>
  <div id=main-content>
    <h1>This is my lovely web page</h1>
    <p>It has lots of lovely content. It has some
      blockquote:</p>
```

Bottom Screenshot (Unquoted Attributes):

```
<title>My lovely web page</title>
</head>

<body id="splash-page">
  <div id="main-content">
    <h1>This is my lovely web page</h1>
    <p>It has lots of lovely content. It has some
      blockquote:</p>
```

In the top screenshot, the attribute values 'splash-page' and 'main-content' are colored blue, while the quoted attribute values 'id=splash-page' and 'id=main-content' are colored black. In the bottom screenshot, the attribute values 'splash-page' and 'main-content' are colored green, while the quoted attribute values 'id="splash-page"' and 'id="main-content"' are colored black.

You may add comments in your HTML, perhaps to make it clear where sections start or end, or to provide a note to remind yourself why you approached the creation of a page in a certain way. What you use comments for isn't important, but the way that you craft a comment is important. The HTML comment looks like this: <!-- this is a comment -->. It's derived from SGML, which starts with an <! and ends with an >; the actual comment is, in effect, inside the opening -- and the closing -- parts. These hyphens tell the browser when to start ignoring text content, and when to start paying attention again. The fact that the double hyphen -- characters signify the beginning and end of the comment means that you should not use double hyphens anywhere inside a comment, even if you believe that your usage of these characters conforms to SGML rules. Single hyphens are allowed, however.

The markup below shows examples of good and bad HTML comments—see the remark associated with each example for more information:

<p>Take the next right.<!-- Look out for the

signpost for 'Castle' --></p> a valid comment

<p>Take the next right.<!-- Look out for -- Castle --></p>

not a valid comment; the double dashes in the middle could be misinterpreted as the end of the comment

<p>Take the next right.<!-- Look out for --- Castle --></p>

a valid comment; 'Look out for' is one comment, 'Castle' is another

<p>Take the next right.

<!-------

This is just asking for trouble. Too

many hyphens! --></p>

a valid comment; don't use hyphens or <> characters to format comment text

<p <!-- class="lively" -->>Wowzers!</p>

It's not possible to comment out attributes inside an HTML element

1.12 Standard XHTML document structure

An XHTML document consists of three main parts:

- DOCTYPE
- Head

- Body

The basic document structure is:

```
<!DOCTYPE ...>
<html ... >
<head> ... </head>
<body> ... </body>
</html>
```

The `<head>` area contains information about the document, such as ownership, copyright, and keywords; and the `<body>` area contains the content of the document to be displayed.

Listing 1 shows you how this structure might be used in practice:

Listing 1. An XHTML example

1. `<?xml version="1.0"?>`
2. `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "DTD/xhtml1-transitional.dtd">`
3. `<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">`
4. `<head>`
`<title>My XHTML Sample Page</title>`
`</head>`
5. `<body bgcolor="white">`
`<center><h1>Welcome to XHTML !</h1></center>`
`</body>`
6. `</html>`

Line 1: Since XHTML is HTML expressed in an XML document, it must include the initial XML declaration `<?xml version="1.0"?>` at the top of the document.

Line 2: XHTML documents must be identified by one of three standard sets of rules. These rules are stored in a separate document called a Document Type Declaration (DTD), and are utilized to validate the accuracy of the XHTML document structure. The purpose of a DTD is to describe, in precCSE terms, the language and syntax allowed in XHTML.

Line 3: The second tag in an XHTML document must include the opening <html> tag with the XML namespace identified by the xmlns=http://www.w3.org/1999/xhtml attribute. The XML namespace identifies the range of tags used by the XHTML document. It is used to ensure that names used by one DTD don't conflict with user-defined tags or tags defined in other DTDs.

Line 4: XHTML documents must include a full header area. This area contains the opening <head> tag and the title tags (<title></title>), and is then completed with the closing </head> tag.

Line 5: XHTML documents must include opening and closing <body></body> tags. Within these tags you can place your traditional HTML coding tags. To be XHTML conformant, the coding of these tags must be well-formed.

Line 6: Finally, the XHTML document is completed with the closing </html> tag.

1.13 Basic text markup

A **markup language** is a modern system for annotating a text in a way that is syntactically distinguishable from that text. The idea and terminology evolved from the "marking up" of manuscripts, i.e. the revision instructions by editors, traditionally written with a blue pencil on authors' manuscripts. Examples are typesetting instructions such as those found in troff and LaTeX, and structural markers such as XML tags. Markup is typically omitted from the version of the text which is displayed for end-user consumption. Some markup languages, like HTML have presentation semantics, meaning their specification prescribes how the structured data is to be presented, but other markup languages, like XML, have no predefined semantics.

A well-known example of a markup language in widespread use today is HyperText Markup Language (HTML), one of the document formats of the World Wide Web. HTML is mostly an instance of SGML (though, strictly, it does not comply with all the rules of SGML) and follows many of the markup conventions used in the publishing industry in the communication of printed work between authors, editors, and printers.

UNIT - 2

XHTML – 2

HTML is defined using the Standard Generalized Markup Language (SGML) which is an ISO standard notation for describing text formatting languages. The Original version of HTML was designed in conjunction with the structure of the web and the first browser. It was designed to specify document structure at a higher and more abstract level. It also defines details of text such as font style, size and color. The use of web became more popular when the first browser MOSAIC was developed and marketed by Netscape. The next browser was Internet explorer by Microsoft.

History of HTML

HTML was the standard developed by Microsoft for web technologies. First HTML standard was HTML 2.0 was released in 1995. Next HTML standard was HTML 3.2 was released in 1997. The latest version of HTML standard is HTML 4.01 was released and approved by W3C in 1999. The XHTML1.0 standard was approved in early 2000 which is a redefinition of HTML 4.01 using XML IE7 and FIREFOX2 (FX2) support XHTML 1.1

2.1 Images

Images can be included to enhance appearance of the document. Most common methods of representing the images are GIF (graphical interchange format and JPEG (joint photographic experts group) format. The former is 8bit color representation whereas the latter is 24 bit color representation. The image tag specifies an image that appears in a document. It has attributes like src which specifies the source of the image.

```
<html xmlns = "http://www.w3.org/1999/xhtml">
<head><title> Images</title></head>
<body>
<h1> Twinkle twinkle
<h2>little star
<h3> how I wonder
<h4>what you are ???
up above the world so high
```

like a diamond in the sky.

```

```

```
</body>
```

```
</HTML>
```

XHTML validation

2.2 Hypertext Links

A hypertext link in a XHTML document acts as a pointer to some resource. It could be an XHTML document anywhere in the web or another place in the document currently displayed. Links that point to another document specifies the address of the document. It could be a filename, complete url, directory path and a filename. Links are specified in an attribute of an anchor tag [which is an inline tag. The anchor tag is the source of a link whereas the document is the target of the link. Links facilitate reader to click on links to learn more about a particular subtopic of interest and also return back to the location of the link](#)

If target is in the same document as the link it is specified in the href attribute value by preceding the id with a pound sign(#)

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

```
<head><title> Images</title></head>
```

```
<body>
```

```
<h1> Twinkle twinkle
```

```
<h2>little star
```

```
<h3> how I wonder
```

```
<h4>what you are ???
```

up above the world so high like a diamond in the sky.

```
<p>
```

```
<a href = "C:\Documents and Settings\Administrator\My Documents\XHTML  
programs\1my.html">The blue hill image document
```

```
</a>
```

```
<p>
```

```
</body>
</HTML>

<html xmlns = "http://www.w3.org/1999/xhtml">
<head><title> Images</title></head>
<body>
<h1 id = "twinkle" > Twinkle twinkle </h1>
<h2>little star</h2>
<h3> how I wonder </h3>
<h4>what you are ???
up above the world so high
like a diamond in the sky.</h4>
<a href = "#twinkle">Which poem
</a>
</body>
</HTML>
```

Twinkle twinkle

little star

how I wonder

what you are ??? up above the world so high like a diamond in the sky.

[Which poem](#)

2.3 Lists

XHTML provides simple and effective ways to specify both ordered and unordered lists `` `` are tags for unordered and ordered lists. Each item in a list is specified with an `` tag. Any tags can appear in a list item including nested lists.

Definition lists

They are used to specify list of terms and their definitions such as glossaries. They are given by `<dl>` tag which is a block tag. The definitions are specified as the content of `<dd>` tag and the definition list is given as the content of a `<dt>` tag

```
<html xmlns = "http://www.w3.org/1999/xhtml">
<head><title> ordered lists</title></head>
<body>
<h3 > Lists of poems</h1>
<ul>
<li> Twinkle twinkle</li>
<li> Baa Baa black sheep</li>
<li> pussy cat </li>
<li> Humpty dumpty</li>
</ul>
</body>
</html>
<body>
<h3 > Lists of poems</h1>
<ol>
<li> Twinkle twinkle</li>
<li> Baa Baa black sheep</li>
<li> pussy cat </li>
<li> Humpty dumpty</li>
</ol>
</body>
```

Lists of poems

1. Twinkle twinkle
2. Baa Baa black sheep
3. pussy cat
4. Humpty dumpty

Lists of poems

- Twinkle twinkle
- Baa Baa black sheep
- pussy cat
- Humpty dumpty

```
head><title> ordered lists</title></head>
<body>
<ol>
<li > Lists of poems
<ol>
<li> Twinkle twinkle</li>
<li> Baa Baa black sheep</li>
<li> pussy cat </li>
<li> Humpty dumpty</li>
</ol>
<ol>
</li>
<li > Lists of stories
<ol>
<li> Thirsty crow</li>
<li> Lion and the mouse</li>
<li> pussy cat </li>
<li> Midas touch</li>
</ol>
</li>
</ol>
```

1. Lists of poems
 1. Twinkle twinkle
 2. Baa Baa black sheep
 3. pussy cat
 4. Humpty dumpty
2. Lists of stories
 1. Thirsty crow
 2. Lion and the mouse
 3. pussy cat
 4. Midas touch

```
<html xmlns = "http://www.w3.org/1999/xhtml">
<head><title> ordered lists</title></head>
<body>
<h3 > Lists of poems </h3>
<dl>
<dt> 111 </dt>
<dd> Twinkle twinkle</dd>
<dt> 222 </dt>
<dd> Pily poly</dd>
<dt> 333 </dt>
<dd> Rudolf the reindeer</dd>
</dl>
</body>
</HTML>
```

Lists of poems

```
111
    Twinkle twinkle
222
    Pily poly
333
    Rudolf the reindeer
```

2.4 Tables

A table is a matrix of cells, each possibly having content. The cells can include almost any element some cells have row or column labels and some have data. A table is specified as the content of a `<table>` tag which is a block tag. A border attribute in the `<table>` tag specifies a border between the cells. Rule specifies the lines that separate the cells.

If border is set to `-border|`, the browser's default width border is used. The border attribute can be set to a number, which will be the border width in pixels(0 is no border no rules). Without the border attribute, the table will have no lines!. Tables are given titles with the `<caption>` tag, which can immediately follow `<table>`.

Each row of a table is specified as the content of a `<tr>` tag. The row headings are specified as the content of a `<th>` tag. The contents of a data cell is specified as the content of a `<td>` tag. The empty cell is specified with a table header tag that includes no content `<th> </th>`.

```
<body>
<table border = "border">
<caption> fruit juice drinks </caption>
<tr>
<th> </th>
<th> Apple </th>
<th> Mango </th>
<th> Strawberry </th>
</tr>
<tr> <th> Breakfast </th>
<th> 0</th>
<th> 1 </th>
<th> 0</th>
</tr>
<tr> <th> lunch </th>
<th> 1</th>
<th> 1 </th>
<th>1 </th>
</tr>
```

```
<tr> <th> dinner </th>
<th> 0 </th>
<th> 1 </th>
<th> 0 </th></tr><table></body></HTML>
```

Colspan Rowspan attributes

A table can have two levels column labels and also row labels. If so, the colspan attribute must be set in the `<th>` tag to specify that the label must span some number of columns.

```
<tr>
<th colspan = "3"> Fruit Juice Drinks </th>
</tr>
<tr>
<th> Orange </th>
<th> Apple </th>
<th> Screwdriver </th>
</tr>
<caption> fruit juice drinks </caption>
<tr>
<td rowspan="2"></td>
<th colspan="3">Juices chart</th>
</tr>
<tr>
<th> </th>
```

If the rows have labels and there is a spanning column label, the upper left corner must be made larger, using rowspan.

```
<table border = "border">
<tr>
<td rowspan = "2"> </td>
<th colspan = "3"> Fruit Juice Drinks
</th>
</tr>
<tr>
<th> Apple </th>
<th> Orange </th>
<th> Screwdriver </th>
```

```
</tr>
...
</table>
```

	fruit	juice	drinks
	Apple	Mango	Strawberry
Breakfast	0	1	0
lunch	1	1	1
dinner	0	1	0

The align attribute controls the horizontal placement of the contents in a table cell. Values are left, right, and center (default) align is an attribute of <tr>, <th>, and <td> elements. The valign attribute controls the vertical placement of the contents of a table cell. Values are top, bottom, and center (default) valign is an attribute of <th> and <td> elements. The cellspacing attribute of <table> is used to specify the distance between cells in a table. The cellpadding attribute of <table> is used to specify the spacing between the content of a cell and the inner walls of the cell.

```
<head><title> simple table</title></head>
<body>
<table border = "border">
<caption> align</caption>
<tr align ="center">
<th> </th>
<th> Apple </th>
<th> Mango </th>
<th> Strawberry </th>
</tr>
<tr>
<th> align </th>
<td align="left">left</td>
<td align="center">center</td>
<td align="right">right</td>
```

```

</tr>
<tr>
<th> valign<br /><br /></th>
<td >default</td>
<td valign="top">top</td>
<td valign="bottom">bottom</td>
</tr>
<table>
</body>
</HTML>

```

align			
	Apple	Mango	Strawberry
align	left	center	right
valign	default	top	bottom

2.5 Forms

A form is the usual way information is gotten from a Web browser to a server. HTML has tags to create a collection of objects that implement this information gathering. The objects are called widgets (e.g., radio buttons and checkboxes). All control tags are inline tags. These controls gather information used from user in the form of either text or button selections. Each control has a value given through the user input. Collectively values of all these controls in a form are called the form data. When the Submit button of a form is clicked, the form's values are sent to the Web server for processing.

The **<form>** tag

All of the widgets, or components of a form are defined in the content of a **<form>** tag which is a block tag. This tag can have many attributes of which the most required attribute is the action. The action attribute specifies the URL of the application on the web server that is to be called when the user clicks the Submit button.

Eg: action = <http://www.cs.ucp.edu/cgi-bin/survey.pl>

If the form has no action, the value of action is the empty string. The method attribute of <form> specifies one of the two techniques, get or post, used to pass the form data to the server. get is the default, so if no method attribute is given in the <form> tag, get will be used. The alternative technique is post. With these techniques the form data is encoded into text string on click of submit button Widgets or Controls Many commonly used controls are created with the <input> tag which specifies the kind of control. It is used for the text, passwords, checkboxes, radio buttons and the action buttons Reset and Submit.

The type attribute of <input> specifies the kind of widget being created. Except Reset and Submit all other controls have a name attribute other than type attribute.

Text: Creates a horizontal box for text input

Default size is 20; it can be changed with the size Attribute. If more characters are entered than will fit, the box is scrolled (shifted) left. If you don't want to allow the user to type more characters than will fit, set max length, which causes excess input to be ignored

```
<input type = "text" name = "Phone" size = "12" >
```

If the contents of the textbox should not be displayed when user types it than a password control should be used. Labeling a text box can be done by adding label control to the text box. Both the controls can be encapsulated. This has several advantages. The text content of the label will indicate content of text box and when a label is selected the cursor is implicitly moved to the control.

```
<form action = "">  
<p>  
<input type = "text" name = "Phone" size = "12" /></p>  
<p>  
<input type = "password" name = "myPassword" size = "12" maxlength="12" />  
</p>  
<p>  
<label>Phone:<input type = "text" name = "Phone" size = "12" /></label>  
</p>
```

Checkboxes

Checkboxes collect multiple choice input. Every checkbox requires a value attribute, which is the widget's value in the form data when the checkbox is _checked. A checkbox that is not _checked contributes no value to the form data. By default, no checkbox is initially _checked. To initialize a checkbox to _checked, the checked attribute must be set to

"checked".

Radio Buttons

Radio buttons are collections of checkboxes in which only one button can be checked at a time. Every button in a radio button group MUST have the same name. If no button in a radio button group is pressed, the browser often presses the first one. Checkboxes and radio buttons are both multiple choice input from the user

```
<form action = "">  
<p>  
<input type = "checkbox" name ="groceries" value = "milk" checked = "checked">  
Milk  
<input type = "checkbox" name ="groceries" value = "bread">  
Bread  
<input type = "checkbox" name = "groceries" value= "eggs">  
Eggs  
</p>  
</form>  
<p><input type = "radio" name = "age" value = "under20" checked = "checked">  
0-19  
<input type = "radio" name = "age" value = "20-35">  
20-35  
<input type = "radio" name = "age" value = "36-50">  
36-50  
<input type = "radio" name = "age" value = "over50">  
Over 50
```

Milk Bread Eggs

0-19 20-35 36-50 Over 50

Menus

Each item of a menu is specified with an `<option>` tag, whose pure text content (no tags) is the value of the item. An `<option>` tag can include the `selected` attribute, which when assigned "selected" specifies that the item is preselected. Menus - created with `<select>` tags. There are two kinds of menus, those that behave like checkboxes and those that behave like

radio buttons (the default). Menus that behave like checkboxes are specified by including the multiple attribute, which must be set to "multiple". The name attribute of <select> is required. The size attribute of <select> can be included to specify the number of menu items to be displayed (the default is 1). If size is set to > 1 or if multiple is specified, the menu is displayed as a pop-up menu

Text areas - created with <textarea>. Usually include the rows and cols attributes to specify the size of the text area. Default text can be included as the content of <textarea>. Scrolling is implicit if the area is overfilled.

Reset and Submit buttons both are created with <input>.

```
<input type = "reset" value = "Reset Form">  
<input type = "submit" value = "Submit Form">
```

Submit has two actions:

1. Encode the data of the form
2. Request that the server execute the server-resident program specified as the value of the action attribute of <form>.

A Submit button is required in every form

```
<form action = ""> <p>
```

With size = 1 (the default)

```
<select name = "groceries">  
  <option> milk </option>  
  <option> bread </option>  
  <option> eggs </option>  
  <option> cheese </option>  
</select>  
</p>
```

```
<textarea name = "aspirations" rows = "3" cols = "40">
```

(Be brief in expressing your views)

```
</textarea>  
</p>  
<p>  
  <input type = "Submit" value = "Submit Form" />  
  <input type = "reset" value = "Reset Form" />
```

```
</p>
```

```
</form>
```

2.6 Frames

Frames are rectangular sections of the display window, each of which can display a different document. Because frames are no longer part of XHTML, you cannot validate a document that includes frames. The `<frameset>` tag specifies the number of frames and their layout in the window. `<frameset>` takes the place of `<body>`. We Cannot have both! `<frameset>` must have either a `rows` attribute or a `cols` attribute, or both (usually the case). Default is 1. The possible values for `rows` and `cols` are numbers, percentages, and asterisks.

A number value specifies the row height in pixels. A percentage specifies the percentage of total window height for the row. An asterisk after some other specification gives the remainder of the height of the window

Examples:

```
<frameset rows = "150, 200, 300">
<frameset rows = "25%, 50%, 25%">
<frameset rows = "50%, 20%, *" >
<frameset rows = "50%, 25%, 25%">
    cols = "40%, *">
```

The `<frame>` tag specifies the content of a frame. The first `<frame>` tag in a `<frameset>` specifies the content of the first frame, etc. An asterisk after some other specification gives the remainder of the height of the window

Examples:

```
<frameset rows = "150, 200, 300">
<frameset rows = "25%, 50%, 25%">
<frameset rows = "50%, 20%, *" >
<frameset rows = "50%, 25%, 25%">
    cols = "40%, *">
```

The `<frame>` tag specifies the content of a frame. The first `<frame>` tag in a `<frameset>` specifies the content of the first frame, etc.

Row-major order is used. Frame content is specified with the `src` attribute. Without a `src` attribute, the frame will be empty (such a frame CANNOT be filled later). If `<frameset>` has fewer `<frame>` tags than frames, the extra frames are empty. Scrollbars are implicitly included if needed (they are needed if the specified document will not fit). If a

name attribute is included, the content of the frame can be changed later (by selection of a link in some other frame) Nested frames - to divide the screen in more interesting ways

```
<head> <title> Nested frames </title> </head>
<frameset cols = "40%, *" >
<frameset rows = "50%, *" >
<frame src = "1my.html" />
<frame src = "12list.html" />
</frameset>
<frameset rows = "20%,35%, *" >
<frame src = "1my.html" />
<frame src = "12list.html" />
<frame src = "2twinkle.html" />
<frame src = "5head.html" />
</frameset>
</frameset>
</html>
```

2.7 Syntactic differences between HTML and XHTML

Case sensitivity: In HTML, tag and attribute names are case insensitive meaning that <FORM>, <form>, and <Form> are equivalent. In XHTML, all tag and attribute names must be in lowercase.

Closing tags: In HTML, closing tag may be omitted if the processing agent can infer their presence. For example, in HTML, paragraph elements often do not have closing tags. For example

<p> During Spring, flowers are born. ...

<p>

During Fall, flowers die.....

HTML documents are case insensitive whereas XHTML documents have to be in lowercase. Closing tags may be omitted in HTML but not in XHTML where all elements must have closing tags except for content tags where it is not a must

<input type= -text|| name=-address|| />

Quoted attribute values : all attribute values must be quoted whether it is numeric or character based

Explicit attribute values: In HTML some attributes are implicit

Quoted attribute values: In HTML, attribute values must be quoted only if there are embedded special characters or white space characters.

Explicit attribute values: In HTML, some attribute values are implicit, that is, they need not be explicitly stated. For example, if the border attribute appears in a <table> tag without a value, it specifies a default width border on the table. For example: <table border>. id and name attributes. HTML markup often uses the name attribute for elements. Element nesting: Although HTML has rules against improper nesting of elements, they are not enforced.

Examples of nesting rules are:

1. An anchor element cannot contain another anchor element, and a form element cannot contain another form element.
2. If one element appears inside another element its closing tag should appear before the closing tag of the outer element.
3. Block elements cannot be nested inside inline elements.
4. Text cannot be nested directly in body or form elements.
5. List elements cannot be directly nested in list elements.

CSS

2.1 Introduction

The CSS1 specification was developed in 1996 by W3C. CSS2 was released in 1998 which added many properties and values to CSS1. CSS3 has been under development since the late 1990s. CSSs provide the means to control and change presentation of HTML documents. CSS is not technically HTML, but can be embedded in HTML documents.

Cascading style sheets were introduced to provide a uniform and consistent way to specify presentation details in XHTML documents. Most of the style tags and attributes are those deprecated from HTML 4.0 in favor of style sheets. Idea of style sheets is not a new concept as it existed in desktop publishing systems and word processors. Style sheets allow you to impose a standard style on a whole document, or even a whole collection of documents. Style sheets can be defined at three levels to specify the style of a document .Hence called Cascading style sheets. Style is specified for a tag by the values of its properties.

For example:

```
<h2 { font-size: 22pt; } >
```

2.2 Levels of Style Sheets

There are three levels of style sheets, in order from lowest level to highest level, are inline, document level, and external. Inline style sheets are specified for a specific occurrence of a tag and apply only to the content of that tag. This application of style, which defeats the purpose of style sheets – that of imposing uniform style on the tags of at least one whole document. Another disadvantage of inline style sheets is that they result in style information, which is expressed in a language distinct from XHTML markup, being embedded in various places in documents.

Document-level style specifications appear in the document head section and apply to the whole body of the document. External style sheets are not part of the documents to which they apply.

They are stored separately and are referenced in all documents that use them. They are written as text files with MIME type text/css.

```
<html xmlns = "http://www.w3.org/1999/xhtml">  
<head>  
  <title> Our first document </title>  
  <style>  
    h2 { font-size: 32pt; font-weight: bold; font-family: 'Times New Roman'; }  
    h3,h4 { font-size: 18pt; font-family: 'Courier New'; font-style: italic;  
            font-weight: bold }  
  </style>  
</head> <body>  
  <h1> Twinkle twinkle little star</h1>  
  <h2> how I wonder </h2>  
  <h3> what you are ???</h3>  
  <h4> up above the world so high</h4>  
  <h5> like a diamond</h5>  
  <h6> in the sky.</h6>  
</body>  
</html>
```

Twinkle twinkle little star

how I wonder

what you are ???

up above the world so high

like a diamond

in the sky.

Linking an External Style sheet

A <link> tag is used to specify that the browser is to fetch and use an external style sheet

file through href. The href attribute of <link> is used to specify the URL of the style sheet document, as in the following example:

```
<link rel = "stylesheet" type = "text/css"
      href = "http://www.wherever.org/termpaper.css"></link>
```

This link must appear in the head of the document. External style sheets can be validated, with the service <http://jigsaw.w3.org/css-validator/> validator-upload.html. External style sheets can be added using other alternate style specification known as file import

- @import url (filename);

Filename is not quoted. Import appears only at the beginning of the content of a style element. The file imported can contain both markup as well as style rules

2.3 Style Specification Formats

The format of a style specification depends on the level of the style sheet. Inline style sheet appears as the value of the style attribute of the tag. The general form of which is as follows:

```
style = "property_1: value_1;
        property_2: value_2;
        ...
        property_n: value_n;"
```

Format for Document-level

Document style specifications appear as the content of a style element within the header of a document, although the format of the specification is quite different from that of inline style sheets. The <style> tag must include the type attribute, set to "text/css—(as there are other style sheets in JavaScript).

The list of rules must be placed in a comment, because CSS is not XHTML. Style element must be placed within the header of a document. Comments in the rule list must have a different form use C comments /*...*/. The general form of the content of a style element is as follows:

```
<style type = "text/css">
/*
rule list/* styles for paragraphs and other tags*/
</style>
```

Form of the rules:

Each style in a rule list has two parts: selector, which indicates the tag or tags affected by the rules. Each property/value pair has the form->property: value

```
Selector { property_1: value_1; property_2: value_2; ... property_n: value_n; }
```

Pairs are separated by semicolons, just as in the value of a <style> tag.

2.4 Selector Forms

Selector can have variety of forms like:

1. Simple selector form
2. Class selector
3. Generic selector
4. Id selector
5. Universal selector
6. Pseudo classes

Simple selector form

Simple selector form is a list of style rules, as in the content of a <style> tag for document-level style sheets. The selector is a tag name or a list of tag names, separated by commas. Consider the following examples, in which the property is font-size and the property value is a number of points :

```
h1, h3 { font-size: 24pt; }
```

```
h2 { font-size: 20pt; }
```

Selectors can also specify that the style should apply only to elements in certain positions in the document .This is done by listing the element hierarchy in the selector.

- Contextual selectors: Selectors can also specify that the style should apply only to elements in certain positions in the document .
- In the eg selector applies its style to the content of emphasis elements that are descendants of bold elements in the body of the document. `body b em {font-size: 24pt;}`
Also called as descendant selectors. It will not apply to emphasis element not descendant of bold face element.

Class Selectors

Used to allow different occurrences of the same tag to use different style specifications.

A style class has a name, which is attached to the tag's name with a period.

```
p.narrow {property-value list}
```

```
p.wide {property-value list}
```

The class you want on a particular occurrence of a tag is specified with the class attribute of the tag.

For example,

```
<p class = "narrow">
```

Once upon a time there lived a king in the place called Ayodhya.

```
</p>
```

...

```
<p class = "wide">
```

Once upon a time there lived a king in the place called Ayodhya.

```
</p>
```

Generic Selectors

A generic class can be defined if you want a style to apply to more than one kind of tag.

A generic class must be named, and the name must begin with a period without a tag name in its name.

For Example:

```
.really-big { ... }
```

Use it as if it were a normal style class

```
<h1 class = "really-big"> This Tuesday is a holiday </h1>...
```

```
<p class = "really-big"> ... </p>
```

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

```
<head>
```

```
<title> Absolute positioning </title>
```

```
<style type = "text/css">
```

```
.regtext {font-family: Times; font-size: 14pt; width: 600px}
```

```
.abstext {position: absolute; top: 25px; left: 50px; font-family: Times; font-size: 24pt; fontstyle:
```

```
italic; letter-spacing: 1em; color: rgb(102,102,102); width: 500px }
```

```
</style>
```

```
</head>
```

```
<body>
```

```
<p class = "regtext">
```

Apple is the common name for any tree of the genus *Malus*, of the family Rosaceae. Apple trees grow in any of the temperate areas of the world. Some apple blossoms are white, but most have stripes or tints of rose. Some apple blossoms are bright red. Apples have a firm and fleshy structure that grows from the blossom. The colors of apples range from green to very dark red. The wood of apple trees is fine-grained and hard. It is, therefore, good for furniture construction. Apple trees have been grown for many centuries. They are propagated by grafting because they do not reproduce themselves.

</p>

<p class = "abstext"> APPLES ARE GOOD FOR YOU </p>

</body>

</html>

Apple is the common name for any tree of the genus *Malus*, of the family Rosaceae. Apple trees grow in any of the temperate areas of the world. Some apple blossoms are white, but most have stripes or tints of rose. Some apple blossoms are bright red. Apples have a firm and fleshy structure that grows from the blossom. The colors of apples range from green to very dark red. The wood of apple trees is fine-grained and hard. It is, therefore, good for furniture construction. Apple trees have been grown for many centuries. They are propagated by grafting because they do not reproduce themselves.

Id Selectors

An id selector allow the application of a style to one specific element. The general form of an id selector is as follows :

#specific-id {property-value list}

Example:

#section14 {font-size: 20} specifies a font size of 20 points to the element

<h2 id =“section14”> Alice in wonderland</h2>

Universal selector

The universal selector, denoted by an asterisk(*), which applies style to all elements in

the document. For example:

```
{color: red;}
```

makes all elements in the document red.

Twinkle twinkle little star

how I wonder

what you are ???

up above the world so high

like a diamond

in the sky.

Pseudo Classes

Pseudo classes are styles that apply when something happens, rather than because the target element simply exists. Names of pseudo classes begin with colons hover classes apply when the mouse cursor is over the element focus classes apply when an element has focus i.e. the mouse cursor is over the element and the left mouse button is clicked.

These two pseudo classes are supported by FX2 but IE7 supports only hover.

```
<html xmlns = "http://www.w3.org/1999/xhtml">
<head> <title> Checkboxes </title>
<style type = "text/css">
input:hover {color: red;}
input:focus {color: green;}
</style> </head> <body>
<form action = ""> <p>
Your name:
<input type = "text" />
```

```
</p> </form>  
</body>  
</html>
```

Your name:

Your name:

Your name:

2.5 Property Values Forms

CSS1 includes 60 different properties in 7 categories(list can be found in W3C website) Fonts, Lists, Alignment of text, Margins, Colors, Backgrounds, Borders. Keywords property values are used when there are only a few possible values and they are predefined

Eg: small, large, medium.

Keyword values are not case sensitive, so Small, SmALL, and SMALL are all the same as small. Number values can be integer or sequence of digits with decimal points and a + or – sign. Length value are specified as number values that are followed immediately by a two character abbreviation of a unit name. There can be no space between the number and the unit name. The possible unit names are px for pixels, in for inches, cm for centimeters, mm for millimeters, pt for points, pc for picas (12 points),em for value of current font size in pixels, ex for height of the letter x. No space is allowed between the number and the unit specification e.g., 1.5 in is illegal!.

Percentage - just a number followed immediately by a percent sign: eg: font size set to 85% means new font size will be 85% of the previous font size value.

URL values: URL property values use a form that is slightly different from references to URLs in links. The actual URL, which can be either absolute or relative, is placed in parentheses and preceded by url, as in the following:

url(protocol://server pathname)

No space should be left between URL and the left parenthesis.

Colors : Color name rgb(n1, n2, n3). Hex form: #B0E0E6 stands for powder blue color.

Property values are inherited by all nested tags, unless overridden.

2.6 Font properties

Font-family

The font-family property is used to specify a list of font name. The browser will use the first font in the list that it supports. For example, the following could be specified. font-family: Arial, Helvetica, Courier

Generic fonts: They can be specified as the font family value for example :serif, sans-serif, cursive, fantasy, and monospace (defined in CSS). Browser has a specific font defined for each generic name. If a font name that has more than one word, it should be single-quoted Eg: font-family: ‘Times New Roman’

Font-size

Possible values: a length number or a name, such as smaller, xx-large, medium , large etc.

Different browsers can use different relative value for the font-size.

Font-variant

The default value of the font-variant property is normal, which specifies the usual character font. This property can be set to small-caps to specify small capital characters.

Font-style

The font-style property is most commonly used to specify italic, as in the following example. Eg: font-style: italic

Font-weights

The font-weight property is used to specify the degree of boldness. For example: font-weight: bold

Font Shorthands

If more than one font property is to be specified than the values may be stated in a list as the value of the font property . The browser will determine from the form of the values which properties to assign. For example, consider the following specification : Eg: font: bold 24pt ‘Times New Roman‘ Palatino Helvetica The order which browser follows is last must be font name, second last font size and then the font style, font variant and font weight can be in any order but before the font size and names. Only the font size and the font family are required in the font value list. Below example displays the fonts.html

```

<html xmlns = "http://www.w3.org/1999/xhtml">
<head> <title> Font properties </title>
<style type = "text/css">
p.big {font-size: 14pt;
font-style: italic;
font-family: 'Times New Roman';
}
p.small {font: 10pt bold 'Courier New';}
h2 {font-family: 'Times New Roman';
font-size: 24pt; font-weight: bold}
h3 {font-family: 'Courier New'; font-size: 18pt}
</style>
</head>
<body>
<p class = "big"> Where there is will there is a way. </p>
<p class = "small"> PractCSE makes a man perfect. </p>
<h2> Chapter 1 Introduction </h2>
<h3> 1.1 The Basics of Web programming </h3>
<h4> Book by Robert Sebesta</h4>
</body> </html>

```

Where there is will there is a way.

Practise makes a man perfect.

Chapter 1 Introduction

1 . 1 The Basics of Web programming

Book by Robert Sebesta

Text Decoration

The text-decoration property is used to specify some special features of the text. The available values are line-through, overline, underline, and none, which is the default. Many browsers underline links. Text decoration is not inherited Eg:line-through, overline, underline, none

Letter-spacing – value is any length property value controls amount of space between

characters in text

- Eg: 3px

```
<html xmlns = "http://www.w3.org/1999/xhtml">  
<head> <title> Text decoration </title>  
<style type = "text/css">  
p.through {text-decoration: line-through}  
p.over {text-decoration: overline}  
p.under {text-decoration: underline}  
</style> </head>  
<body>  
<p class = "through">  
Twinkle twinkle little star how i wonder what you are!!!! </p>  
<p class= "over">  
Twinkle twinkle little star how i wonder what you are!!!! </p>  
<p class = "under">  
Twinkle twinkle little star how i wonder what you are!!!! </p>  
</body></html>
```

~~Twinkle twinkle little star how i wonder what you are!!!!~~

Twinkle twinkle little star how i wonder what you are!!!!

Twinkle twinkle little star how i wonder what you are!!!!

```
<html xmlns="http://www.w3.org/1999/xhtml">  
<head>  
<title>CSS Example</title>  
<link rel="stylesheet" type="text/css" href="ch03_eg01.css" />  
</head>  
<body>  
<h1>Simple CSS Example</h1>  
<p>This simple page demonstrates how CSS can be used to control the presentation of an  
XHTML document.</p>
```

```
<p class="important">This paragraph demonstrates the use of the <code>class</code>
attribute.</p>
</body>
</html>

/* CSS Document for ch03_eg01.html */

body {
    font-family: arial, verdana, sans-serif;
    background-color: #efefef;
}

h1 {
    color: #666666;
    font-size: 22pt;
}

p {
    color: #999999;
    font-size: 10pt;
}

p.important {
    border: solid black 1px;
    background-color: #ffffff;
    padding: 5px;
    margin: 15px;
    width: 40em;
}
```

Simple CSS Example

This simple page demonstrates how CSS can be used to control the presentation of an XHTML document.

This paragraph demonstrates the use of the `class` attribute.

```
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title>CSS Example</title>
<link rel="stylesheet" type="text/css" href="ch03_eg02.css" />
</head>
<body>
<p class="one">The first paragraph of text should be displayed in a sans-serif font.</p>
```

```
<p class="two">The second paragraph of text should be displayed in a serif font.</p>
<p class="three">The third paragraph of text should be displayed in a monospaced
font.</p>
<p class="four">The fourth paragraph of text should be displayed in a cursive font.</p>
<p
class="five">The fifth paragraph of text should be displayed in a fantasy font.</p>
</body>
</html>

CSS Document for ch03_eg02.html */
p.one {font-family:arial, verdana, sans-serif;}
p.two {font-family:times, "times new roman", serif;}
p.three {font-family:courier, "courier new", monospace;}
p.four {font-family: Zapf-Chancery, Santivo, cursive;}
p.five {font-family:Cottonwood, Studz, fantasy;}
```

The first paragraph of text should be displayed in a sans-serif font.

The second paragraph of text should be displayed in a serif font.

The third paragraph of text should be displayed in a monospaced font.

The fourth paragraph of text should be displayed in a cursive font.

The fifth paragraph of text should be displayed in a fantasy font.

```
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title>CSS Example</title>
<link rel="stylesheet" type="text/css" href="ch03_eg08.css" /></head>
<body>
<h1>Lengths</h1>
<p class="px">The length used here is 12 px</p> <p class="pt">The length used here is
12
pt</p>
<p class="pc">The length used here is 2 pc</p> <p class="in">The length used here is
0.5in</p>
<p class="cm">The length used here is 1cm</p> <p class="mm">The length used here is
```

```
12mm</p>
<p class="em">The length used here is 1.5em</p> <p class="ex">The length used here is
1.5ex</p>
</body> </html>
/* CSS Document for ch03_eg08.html */
p {font-family:arial; font-size:12pt;}
/* lengths */
p.px {font-size:12px;}
p.pt {font-size:12pt;}
p.pc {font-size:2pc;} p.in
{font-size:0.5in;} p.cm
{font-size:1cm;} p.mm
{font-size:12mm;} p.em
{font-size:1.5em;}
p.ex {font-size:1.5ex;}
```

Lengths

The length used here is 12 px

The length used here is 12 pt

The length used here is 2 pc

The length used here is 0.5in

The length used here is 1cm

The length used here is 12mm

2.7 List properties

It is used to specify style of bullets or sequencing values in list items. The list-style-type of Unordered lists can be set to disc,circle,square or none. Bullet can be a disc (default), a square, or a circle. Set it on either the or tag On , it applies to list items

```
<h3> Some Common Single-Engine Aircraft </h3>
<ul style = "list-style-type: square">
<li> Cessna Skyhawk </li>
<li> Beechcraft Bonanza </li>
<li> Piper Cherokee </li> </ul>
```

On , list-style-type applies to just that item

```
<h3> Some Common Single-Engine Aircraft </h3>
<ul>
<li style = "list-style-type: disc">
Cessna Skyhawk </li>
<li style = "list-style-type: square">
Beechcraft Bonanza </li>
<li style = "list-style-type: circle">
Piper Cherokee </li>
```

Could use an image for the bullets in an unordered list.

Example:<li style = "list-style-image: url(bird.jpg)">

```
<html>
</head><body>
<h3> Name of subjects offered</h3>
<ul style = "list-style-type: square">
<li> web programming</li>
<li> Data structures</li>
<li> Compilers design </li>
</ul>
<h3> Name of subjects offered</h3>
<ul>
<li style = "list-style-type: disc">
web programming </li>
```

```
<li style = "list-style-type: square">  
Data structures</li>  
<li style = "list-style-type: circle">  
Compilers design </li>  
</ul></body>  
</html>
```

Name of subjects offered

- web programming
- Data structures
- Compilers design

Name of subjects offered

- ◆ web programming
- Data structures
- Compilers design

When ordered lists are nested, it is best to use different kinds of sequence values for the different levels of nesting. The list-style-type can be used to change the sequence values. Below table lists the different possibilities defined by CSS1. Property value Sequence type first four values

Decimal Arabic numerals 1, 2, 3, 4

upper-alpha Uc letters A, B, C, D

lower-alpha Lc letters a, b, c, d

upper-roman Uc Roman I, II, III, IV

lower-roman Lc Roman i, ii, iii, iv

CSS2 has more, like lower-greek and hebrew

```
<?xml version = "1.0"?>  
<!DOCTYPE html PUBLIC "-//w3c//DTD XHTML 1.1//EN"  
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">  
<html xmlns = "http://www.w3.org/1999/xhtml">  
<head> <title> Sequence types </title>  
<style type = "text/css">
```

```
ol {list-style-type: upper-roman;}  
ol ol {list-style-type: upper-alpha;}  
ol ol ol {list-style-type: decimal;}  
</style>  
</head><body>  
<h3> Aircraft Types </h3>  
<ol>  
<li> General Aviation (piston-driven engines)</li>  
<ol>  
<li> Single-Engine Aircraft</i>  
<ol>  
<li> Tail wheel </li>  
<li> Tricycle </li>  
</ol>  
</li>  
<li> Dual-Engine Aircraft  
<ol>  
<li> Wing-mounted engines </li>  
<li> Push-pull fuselage-mounted engines </li>  
</ol>  
</li>  
</ol>  
</li>
```

Aircraft Types

- I. General Aviation (piston-driven engines)
 - A. Single-Engine Aircraft
 - 1. Tail wheel
 - 2. Tricycle
 - B. Dual-Engine Aircraft
 - 1. Wing-mounted engines
 - 2. Push-pull fuselage-mounted engines
- II. Commercial Aviation (jet engines)
 - A. Dual-Engine
 - 1. Wing-mounted engines
 - 2. Fuselage-mounted engines
 - B. Tri-Engine
 - 1. Third engine in vertical stabilizer
 - 2. Third engine in fuselage

2.8 Colors

Colors are a problem for the Web for two reasons:

1. Monitors vary widely
2. Browsers vary widely

There are three color collections

1. There is a larger set, the Web Palette 216 colors. Use hex color values of 00, 33, 66, 99, CC, and FF
2. Any one of 16 million different colors due to 24 bit color rep
3. There is a set of 16 colors that are guaranteed to be displayable by all graphical browsers on all color monitors

black 000000 green 008000

silver C0C0C0 lime 00FF00

gray 808080 olive 808000

white FFFFFF yellow FFFF00

maroon 800000 navy 000080

red FF0000 blue 0000FF

purple 800080 teal 008080

fuchsia FF00FF aqua 00FFFF

Color properties

The color property specifies the foreground color of XHTML elements. For example, consider the following small table

```
<style type = "text/css">
th.red {color: red}
th.orange {color: orange}
</style> ...
<table border = "5">
<tr>
<th class = "red"> Apple </th>
<th class = "orange"> Orange </th>
<th class = "orange"> Screwdriver </th>
</tr>
</table>
```

The background-color property specifies the background color of elements.

2.9 Alignment of Text

The text-indent property allows indentation. Takes either a length or a % value. The text-align property has the possible values, left (the default), center, right, or justify.

Sometimes we want text to flow around another element - the float property. The float property has the possible values, left, right, and none (the default). If we have an element we want on the right, with text flowing on its left, we use the default text-align value (left) for the text and the right value for float on the element we want on the right.

```
<html xmlns = "http://www.w3.org/1999/xhtml">  
<head> <title> The float property </title>  
<style type = "text/css">  
img {float: right}
```

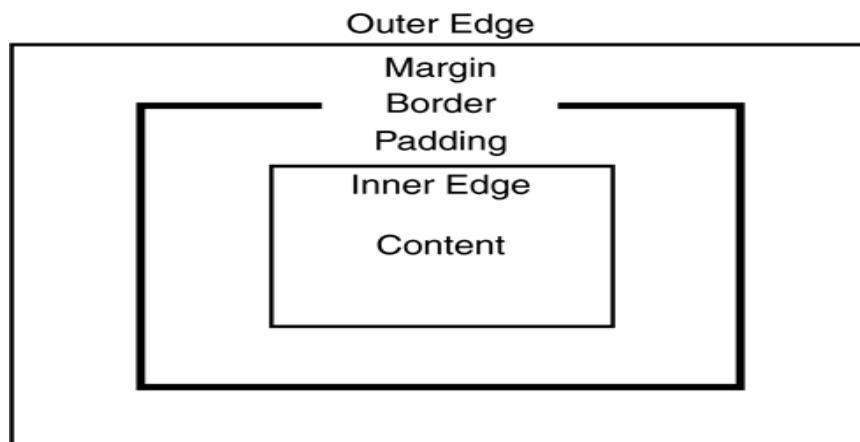
The Box Model

```
</style>  
</head>  
<body> <p>  
<img src = "c210new.jpg" alt = "Picture of a Aircraft" />  
</p> <p>
```

This is a picture of a Cessna 210. The 210 is the flagship single-engine Cessna aircraft. Although the 210 began as a four-place aircraft, it soon acquired a third row of seats, stretching it to a six-place plane.

```
</p> </body>  
</html>
```

2.10 The Box Model



Borders

Every element has a border-style property. It Controls whether the element has a border and if so, the style of the border. The styles of one of the four sides of an element can be set with border-style values: none, dotted, dashed, and double border-width – thin, medium (default), thick, or a length value in pixels. Border width can be specified for any of the four borders (e.g., border-top-width) bordercolor – any color. Border color can be specified for any of the four borders (e.g., border-topcolor)

```
<html xmlns = "http://www.w3.org/1999/xhtml">
<head> <title> Table borders </title>
<style type = "text/css">
table {border-top-width: medium;
border-bottom-width: thick;
border-top-color: red;
border-bottom-color: green;
border-top-style: dotted;
border-bottom-style: dashed;
}
p {border-style: dashed; border-width: thin;
border-color: green
} </style> </head>
<body> <table border = "5">
<caption> Diet chart </caption>
<tr>
<th> </th>
<th> Fruits </th>
<th> vegetables </th>
<th> Carbohydrates </th>
</tr> <tr>
<th> Breakfast </th>
<td> 0 </td>
<td> 1 </td>
<td> 0 </td>
```

Programming the Web

```
</tr> <tr>
<th> Lunch </th>
<td> 1 </td>
<td> 0 </td>
<td> 0 </td> </tr> <tr>
<th> Dinner </th>
<td> 0 </td>
<td> 0 </td>
<td> 1 </td>
</tr>
</table> <p>
```

If you strictly follow the chart you can easily lose weight.

```
</p> </body>
```

```
</html>
```

	Fruits	vegetables	Carbohydrates
Breakfast	0	1	0
Lunch	1	0	0
Dinner	0	0	1



If you strictly follow the chart you can easily lose weight.

Margin

The space between the border of an element and its neighbor element. The margins around an element can be set with margin-left, etc. - just assign them a length value

```
<img src = "c210.jpg " style = "float: right;
margin-left: 0.35in; margin-bottom: 0.35in" />
```

Padding – the distance between the content of an element and its border Controlled by padding, padding-left, etc. bottom, left, or right

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

```
<head><title> Margins and Padding
</title><style type = "text/css">
p.one {margin: 0.2in;
padding: 0.2in;
background-color: #C0C0C0;
border-style: solid;
}
p.two {margin: 0.1in;
padding: 0.3in;
background-color: #C0C0C0;
border-style: solid;
}
p.three {margin: 0.3in; padding: 0.1in;
background-color: #C0C0C0;
border-style: solid; } p.four
{margin:0.4in; background-
color: #C0C0C0;} p.five
{padding: 0.4in; background-
color: #C0C0C0;
}
</style> </head>
</style> </head> <body>
<p> Here is the first line. </p>
<p class = "one">
Style sheets allow you to impose a standard style on a whole document, or even a whole
collection of documents <br /> [margin = 0.2in, padding = 0.2in]
</p>
<p class = "two">
Style sheets allow you to impose a standard style on a whole document, or even a whole
collection of documents. <br /> [margin = 0.1in,
padding = 0.3in]
</p>
<p class = "three">
```

Style sheets allow you to impose a standard style on a whole document, or even a whole collection of documents
 [margin = 0.3in, padding = 0.1in]

Style sheets allow you to impose a standard style on a whole document, or even a whole collection of documents
[margin = 0.2in, padding = 0.2in]

Style sheets allow you to impose a standard style on a whole document, or even a whole collection of documents.
[margin = 0.1in, padding = 0.3in]

Style sheets allow you to impose a standard style on a whole document, or even a whole collection of documents
[margin = 0.3in, padding = 0.1in]

Style sheets allow you to impose a standard style on a whole document, or even a whole collection of documents
[margin = 0.4in, no padding, no border]

Style sheets allow you to impose a standard style on a whole document, or even a whole collection of documents
[padding = 0.4in, no margin, no border]

This is my last session.

2.11 Background Images

The background-image property is used to place an image in the background of an element. Repetition can be controlled. Background image can be replicated to fill the area of the element. This is known as tiling. background-repeat property possible values: repeat (default), no-repeat, repeat-x, or repeat-y. background-position property. Possible values: top, center, bottom, left, or right.

```
<html xmlns = "http://www.w3.org/1999/xhtml">  
<head> <title> Background images </title>  
<style type = "text/css">  
body {background-image: url(c172.gif);}  
p {margin-left: 30px; margin-right: 30px;  
margin-top: 50px; font-size: 14pt;}  
</style>  
</head>
```

```
<body> <p >
```

The Cessna 172 is the most common general aviation airplane in the world. It is an allmetal, single-engine piston,at sea level is 720 feet per minute.

```
</p> </body>
```

```
</html>
```

This is a picture of a Cessna 210. The 210 is the flagship single-engine Cessna aircraft. Although the 210 began as a four-place aircraft, it soon acquired a third row of seats, stretching it to a six-place plane. The 210 is classified as a high performance airplane, which means its landing gear is retractable and its engine has more than 200 horsepower. In its first model year, which was 1960, the 210 was powered by a 260 horsepower fuel-injected six-cylinder engine that displaced 471 cubic inches. The 210 is the fastest single-engine airplane ever built by Cessna.



2.12 Conflict Resolution

When two or more rules apply to the same tag there are resolutions for deciding which rule applies. In-line style sheets have precedence over document style sheets. Document style sheets have precedence over external style sheets. Within the same level there can be conflicts a tag may be used twice as a selector

```
h3{color:red;} body h3 {color: green;}
```

A tag may inherit a property and also be used as a selector. Style sheets can have different sources:

The browser itself may set some style

eg: In FX2 min font size can be set in Tools-Options-Advanced window

The author of a document may specify styles. The user, through browser settings, may specify styles. Individual properties can be specified as important or normal.

Eg: p.special{font-style: italic !important; font-size :14}

This property is known as weight of a specification. Conflict resolution is a multistage sorting process. The first step in the process is to gather the style specifications from the three possible levels of style sheets. These specifications are sorted into order by the relative precedence of the style sheet levels. This is done according to the following rules, in which the first has the highest precedence. From highest to lowest

1. Important declarations with user origin
2. Important declarations with author origin
3. Normal declarations with author origin
4. Normal declarations with user origin
5. Any declarations with browser (or other user agent) origin Tie-Breakers

Conflict resolution by Specificity (high to low)

1. id selectors
2. Class and pseudo-class selectors
3. Contextual selectors
4. General selectors

Position

Essentially, later has precedence over earlier. Most recently seen specification is the one which gets more precedence. Sorting process to resolve the style specification is known as cascade.

UNIT - 3

JAVASCRIPT

3.1 Overview of Javascript

JavaScript is a sequence of statements to be executed by the browser. It is most popular scripting language on the internet, and works in all major browsers, such as IE, FireFox, chrome, opera safari. Prerequisite –HTML/XHTML

Origins

It is originally known as LiveScript, developed by Netscape. It became a joint venture of Netscape and Sun in 1995, and was renamed as JavaScript. It was standardized by the European computer Manufacturers Association as ECMA-262. ISO-16262. Current standard specifications can be found at

<http://www.ecma-international.org/publications/standardsEcma-262.htm>

Collections of JavaScript code scripts and not programs.

What is JavaScript?

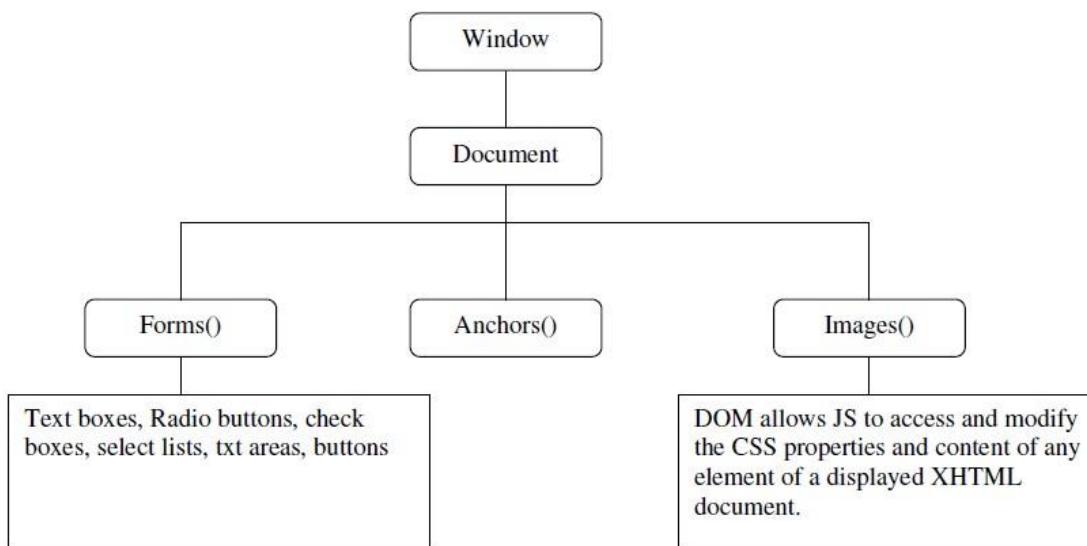
1. JavaScript was designed to add interactivity to HTML pages.
2. JavaScript is a scripting language.
3. A scripting language is a lightweight programming language.
4. It is usually embedded directly into HTML pages.
5. JavaScript is an interpreted language (Scripts are executed without preliminary compilations)

JavaScript can be divided into three parts.

- 1. The Core:** It is a heart of the language, including its operators, expressions, statements and subprograms.
- 2. Client Side:** It is a collection of objects that support control of a browser and interactions with users. Eg. With JavaScript an XHTML document can be made to be responsible to user inputs. Such as mouse clicks and keyboard use.
- 3. Server side:** It is a collection of objects that make the language useful on a Web server. Eg. To support communication with a DBMS. Client side JavaScript is an XHTML embedded scripting language. We refer to every collection of JavaScript code as a script. An XHTML document can include any number of embedded scripts. The HTML

Document Object Model(DOM) is the browsers view of an HTML page as an object hierarchy, starting with the browser window itself and moving deeper into the page, including of the elements on the page and their attribute.

Fig: The HTML DOM



The top level object is window. The document object is a child of window and all the objects that appear on the page are descendants of the document object. These objects can have children of their own. Eg. Form objects generally have several child objects , including textboxes, radio buttons and select menus.

JavaScript and java

Document Forms() Anchors() Images()

Window Text boxes, Radio buttons, check boxes, select lists, txt areas, buttons

DOM allows JS to access and modify the CSS properties and content of any element of a displayed XHTML document.

JavaScript and java is only related through syntax.

JavaScript support for OOP is different from that of Java.

JavaScript is dynamically typed.

Java is strongly typed language. Types are all known at compile time and operand types are checked for compatibility. But variables in JavaScript need not be declared and are dynamically typed, making compile time type checking impossible.

Objects in Java are static -> their collection of data number and methods is fixed at

compile time.

JavaScript objects are dynamic: The number of data members and methods of an object can change during execution.

Uses of JavaScript

Goal of JavaScript is to provide programming capability at both server and the client ends of a Web connection. Client-side JavaScript is embedded in XHTML documents and is interpreted by the browser. This transfer of load from the often overloaded server to the normally under loaded client can obviously benefit all other clients. It cannot replace server side computations like file operations, database access, and networking.

JavaScript can be used as an alternative to Java applets. Java applets are downloaded separately from the XHTML documents that call them but JavaScript are integral part of XHTML document, so no secondary downloading is necessary. Java applets far better for graphics files scripts.

Interactions with users through form elements, such as buttons and menus, can be conveniently described in JavaScript. Because events such as button clicks and mouse movements are easily detected with JavaScript they can be used to trigger computations and provide feedback to the users.

Eg. When user moves the mouse cursor from a textbox, JavaScript can detect that movement and check the appropriateness of the text box's value. Even without forms, user interactions are both possible and simple to program. These interactions which take place in dialog windows include getting input from the user and allowing the user to make choices through buttons. It is also easy to generate new content in the browser display dynamically.

Event driven computation

Event driven computation means that the actions often are executed in response to actions often are executed in response to actions of the users of doc, actions like mouse clicks and form submissions. This type of computation supports user interactions through XHTML form elements on the client display. One of the common uses of JS is client end input data validation values entered by users will be checked before sending them to server for further processing. This becomes more efficient to perform input data checks and carry on this user dialog entirely on the client. This saves both server time and internet time.

Browsers and XHTML/JS documents.

It is an XHTML document does not include embedded scripts, the browser reads the lines of the document and renders its window according to the tags, attributes and content it finds when a JavaScript script is encountered in the doc, the browser uses its JS interpreter to execute the script. When the end of script reached, the browser goes back to reading the XHTML document and displaying its content.

JS scripts can appear in either part of an XHTML document, the head or the body, depending on the purpose of the script. Scripts that produce content only when requested or that react to user interactions are placed in the head of the document. -> Function definition and code associated with form elements such as buttons. Scripts that are to be interpreted just once, when the interpreter finds them are placed in the document body. Accordingly, the interpreter notes the existence of scripts that appear in the head of a document, but it does not interpret them while processing the head. Scripts that are found in the body of a document are interpreted as they are found.

3.2 Object orientation and Javascript

JavaScript is object based language. It doesn't have classes. Its objects serve both as objects and as models of objects. JavaScript does not support class based inheritance as is supported in OO language. CTT-Java. But it supports prototype based inheritance i.e a technique that can be used to simulate some of the aspects of inheritance. JavaScript does not support polymorphism. A polymorphic variable can reference related objects of different classes within the same class hierarchy. A method call through such a polymorphic variable can be dynamically bound to the method in the objects class.

JavaScript Objects

JavaScript objects are collection of prospectus, which corresponds to the members of classes in Java & C++. Each property is either a data property or a function or method property.

1. Data Properties

- a. Primitive Values (Non object Types)
- b. Reference to other objects

2. Method Properties –methods.

Primitives are non object types and are used as they can be implemented directly in hardware resulting in faster operations on their values. These are accessed directly-like scalar types in java & C++ called value types. All objects in a JavaScript programs are directly accessed through variables. Such a variable is like a reference in java. The properties of an object are referenced by attaching the name of the property to the variable that references the object. Eg. If myCar variable referencing an object that has the property engine, the engine property can be referenced with myCar.engine.

The root object in JavaScript is object. It is ancestor through prototype inheritance, of all objects. Object is most generic of all objects, having some methods but no data properties. All other objects are specializations of object, and all inherit its methods.

JavaScript object appears both internally and externally as a list of property/value pairs. Properties are names values are data values of functions. All functions are objects and are referenced through variables. The collection of properties of JavaScript is dynamic – Properties can be added or deleted at any time.

3.3 General syntactic Characteristics

1. JavaScript are embedded either directly or indirectly in XHTML documents.
2. Scripts can appear directly as the content of a <script> tag.
3. The type attribute of <script> must be set to `text/JavaScript`.
4. The JavaScript can be indirectly embedded in an XHTML document using the src attribute of a <script> tag, whose value is name of a file that contains the script.

Eg. `<script type="text/JavaScript" src="tst_number.js"></script>`

Closing tag is required even if script element has src attribute included.

The indirect method of embedding JavaScript in XHTML has advantages of

- 1) Hiding the script from the browser user.
- 2) It also avoids the problem of hiding scripts from older browsers.
- 3) It is good to separate the computation provided by JavaScript from the layout and presentation provided by XHTML and CSS respectively. But it is sometimes not convenient and cumbersome to place all JavaScript code in separate file JavaScript identifiers or names are similar to programming languages.
 1. must begin with (-), or a letter. Subsequent characters may be letters, underscores or digits.
 2. No length limitations for identifiers.

3. Case sensitive
4. No uppercase letters.

Reserved words are break delete function return typeof case do if switch var catch else in this void continue finally instanceof throw while default for new try with

JavaScript has large collection of predefined words

alert

open

java

self

Comments in JavaScript

// - Single line

/* */ -Multiple line

Two issues regarding embedding JavaScript in XHTML documents.

1) There are some browsers still in use that recognize the <script> tag but do not have JS interpreters. These browsers will ignore the contents of the script element and cause no problems.

2) There are still a few browsers in use that are so old they do not recognize <script> tag. These browsers will display the contents of the script elements as if it were just text. Therefore it has been customary to enclose the contents of all script elements in XHTML comments to avoid this problem. XHTML validator also has a problem with embedded JS. When embedded JS happens to include recognizable tags.

For eg
 in output of JS-they often cause validation errors.

Therefore we have to enclose embedded JS in XHTML comments. XHTML comment introduction (<! - -) works as a hiding prelude to JS code. Syntax for closing a comment that encloses JS code is different. It is usual XHTML comment closer but it must be on its own line and preceded by two slashes.

Eg. <!--

-- JS --

//-->

Many more problem are associated with putting embedded JavaScript in comments in XHTML document.

Solution : Put JavaScript scripts of significant style in separate files.

Use of ; in JS is unusual

When EOL coincides with end of statement, the interpreter effectively insects a semicolon there, but this leads to problems.

Eg. return x;

Interpreter puts; after return making x an illegal orphan.

Therefore put JS statements on its own line when possible and terminate each statement with a semicolon. If stmt does not fit in one line, break the stmt at a place that will ensure that the first line does not have the form of a complete statement.

```
<?xml version = -4.0 encoding = -utf-8?>
<!DOCTYPE html PUBLIC -//w3c//DTD XHTML 1.1//EN//
http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd>
<! - -hello.html
```

8

A trivial hello world example of XHTML/JavaScript

```
-->
<html xmlns = -http://www.w3.org/1999/xhtml|.>
<head>
<title> Hello World</title>
</head>
<body>
<script type = -text/javascript|>
<!--
Document.write(-Hello, fellow Web programmers!|);
//-->
</script>
</body>
</html>
```

3.4 Primitives, operations, and expressions

The primitive data types, operations and expressions of JavaScript.

Primitive Types:

Pure primitive types : Number, String, Boolean, Undefined and null. JavaScript includes predefined objects that are closely related to the number, string and Boolean types named number, string and Boolean. These are wrapper objects. Each contains a property that stores a value of the corresponding primitive type. The purpose of the wrapper object is to provide properties and methods that are convenient for use with values of the primitive types.

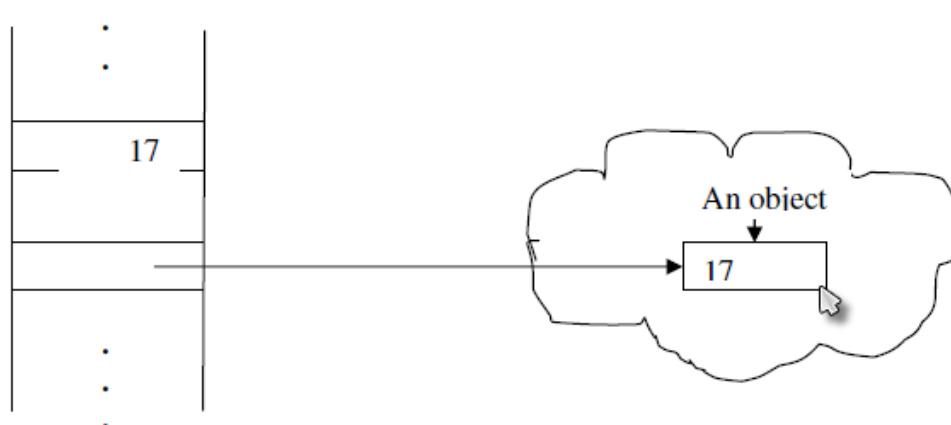
In case of numbers : Properties are more useful.

In case of string : Methods are more useful.

Because JavaScript coerces values between the number type and number objects and between the string type and objects, the methods of number and string can be used on variables of the corresponding primitive types.

Difference between primitives and objects :

Fig:



Prim is a primitive variable with value 17 and obj is a number object whose property value is 17. Fig shows how they are stored.

Numeric and String literals:

All numeric literals are values of type number. The numeric values of JavaScript are represented internally in double precision floating point form, Numeric values in JavaScript are called numbers because of single numeric data type. Literal numbers in a script can have forms of either integers or floating point values. Integer literals are strings of digits.

Floating point literals can have decimal points or exponents or both.

Legal numeric literals: 72, 7.2, .72, 72, 7E2, 7e2, .7e2, 7.e2, 7.2E-2.

Integers in Hexadecimal form 0x or 0X. String Literal: Sequence of 0 or more characters delimited by either single quotes or double quotes. They can include characters specified with escape sequences, such as \n and \t. If you want an actual single quote character in a string literal that is delimited by single quotes, embedded single quote must be preceded by a backslash.

_ You're the most freckly person I've ever met"

-D:\bookfiles\ -> Jo embed\

__or ____> Null string

3.5 Screen output and keyboard input

A JavaScript is interpreted when the browser finds the script in the body of the XHTML document. Thus the normal screen for the JavaScript is the same as the screen in which the content of the host XHTML document is displayed. JS models the XHTML document with the document object. The window in which the browser displays an XHTML document is modeled with the window object. It includes two properties document and window.

Document -> Refers to document object.

Window -> self referential and refers to window object.

Methods -> write

Write is used to create script o/p, which is dynamically created XHTML document content.

Eg. `document.write("The result is : " + result + "
");`

Because write is used to create XHTML code, the only useful punctuations in its parameter is in the form of XHTL tags. Therefore the parameter to write often includes `
` writeln methods implementing adds `\n` to its parameter. As browsers neglects line breaks when displaying XHTML, it has no effect on the output. Window object is JS model for the browser window. It includes three methods that create dialog boxes for three specific kinds of user interactions. The default object for JS is window object currently being displayed, so calls to these methods need not include an object reference.

Alert method opens a dialog window and displays its parameter in that window. It also displays an OK button. The parameter string to alert is not XHTML code, it is plain text.

There fore the string parameter to alert may include \n but never should include
. Confirm method opens a dialog window in which it displays its string parameter, along with two buttons OK and Cancel. Confirm returns a Boolean value that indicates the users button input

True -> for OK

False-> for cancel.

Eg. var question = confirm("Do you want to continue this download?");

After the user responds to one of the button in the confirm dialog window script can test the variable, question and react accordingly. Prompt method creates a dialog window that contains a text box which is used to collect a string of input from the user, which prompt returns as its value. The window also includes two buttons, OK and Cancel, prompt takes two parameters the string that prompts the user for input and a default string in case the user does not type a string before pressing one of the two buttons. In many cases an empty string is used for the default input.

Eg. name name=prompt("What is your name..");

Alert, prompt and confirm cause the browser to wait for a user response.

Alert – OK

Prompt-OK, Cancel

Confirm-OK, Cancel.

Eg.

```
<html>
<head>
</title> roots.html</title>
</head>
<body>
<script type = "text/javascript" src = "roots.js">
</script>
</body>
</html>
//roots.js
// Compute the real roots of a given quadratic equation. If the roots are imaginary,
//this script displays NaN, because that is what results from taking the square root of
```

```

//a negative number.

//Get the coefficients of the equation of the equation from the user
var a = prompt("What is the value of _a'?\n", "-");
var b = prompt("What is the value of _b'?\n", "-");
var c = prompt("What is the value of _c'?\n", "-");
// Compute Square root
var root_part = Math.sqrt(b * b - 4.0 * a * c);
var denom = 2.0 * a;
//Compute and Display
Var root1 = (-b + root_part) / denom; Var root2 = (-b -
root_part) / denom; document.write("The first root is
:",root1, "  
"); document.write("The second root is
:",root2, "  
");

```

3.7 Control Statements

Control expression control the order of execution of statements. Compound statements in JavaScript are syntactic constructs for sequences of statements whose execution they control. Compound statement sequence of statements deleted by braces.

Control construct is a control statement whose execution it controls. Compound statements are not allowed to create local variables.

Control Expressions.

Control statements flow.

Eg. primitive values, relational expression and compound expressions. Result of evaluating a control expression is Boolean value true or false.

For strings.

true - string

false-null string

For number

True- any number

False-0

Operation	Operator
Is equal to	<code>==</code>
Is not equal to	<code>!=</code>
Is less than	<code><</code>
Is greater than	<code>></code>
Is less than or equal to	<code><=</code>
Is greater than or equal to	<code>>=</code>
Is strictly equal to	<code>=====</code>
Is strictly not equal to	<code>!= =</code>

If two operands are not of the same type and operator is neither `==` or `!=`, JS will convert to a single type. Eg. If one is string and other is number, JS will convert string to a number. If one operand is Boolean and other is not, then Boolean is converted to a number.(1 for true, 0 for false)

3.7 Object creation and modification

Create objects with the `new` keyword followed by a constructor. For example, the following program creates a `TwoDPoint` object and prints its fields:

```
class OriginPrinter {
    public static void main(String[] args) {
        TwoDPoint origin; // only declares, does not allocate
        // The constructor allocates and usually initializes the object
        origin = new TwoDPoint();
        // set the fields
        origin.x = 0.0;
        origin.y = 0.0;
        // print the two-d point
        System.out.println(
            "The origin is at " + origin.x + ", " + origin.y);
    } // end main
} // end OriginPrinter
```

The _ ‘ is the member access separator.

A constructor invocation with new is required to allocate an object. There is no C++ like static allocation.

To compile this class, put it in a file called OriginPrinter.java in the same directory as TwoDPoint.java and type:

```
$ javac OriginPrinter.java
```

Multiple Objects

In general there will be more than one object in any given class. Reference variables are used to distinguish between different objects of the same class.

For example, the following program creates two two-d point objects and prints their fields:

```
class TwoPointPrinter {  
    public static void main(String[] args) {  
        TwoDPoint origin; // only declares, does not allocate  
        TwoDPoint one; // only declares, does not allocate  
  
        // The constructor allocates and usually initializes the object  
        origin = new TwoDPoint();  
        one = new TwoDPoint();  
  
        // set the fields  
        origin.x = 0.0;  
        origin.y = 0.0;  
        one.x = 1.0;  
        one.y = 0.0;  
  
        // print the 2D points  
        System.out.println(  
            "The origin is at " + origin.x + ", " + origin.y);  
        System.out.println("One is at " + one.x + ", " + one.y);  
    } // end main  
}  
// end TwoPointPrinter
```

one and origin are two different reference variables pointing to two different point objects. It's not enough to identify a variable as a member of a class like x or y in the example above. You have to specify which object in the class you're referring to.

It is possible for two different reference variables to point to the same object.

When an object is no longer pointed to by any reference variable (including references stored deep inside the runtime or class library) it will be marked for garbage collection.

For example, the following program declares two TwoDPoint reference variables, creates one two-d point object, and assigns that object to both variables. The two variables are equal.

```
class EqualPointPrinter {  
    public static void main(String[] args) {  
  
        TwoDPoint origin1; // only declares, does not allocate  
        TwoDPoint origin2; // only declares, does not allocate  
  
        // The constructor allocates and usually initializes the object  
        origin1 = new TwoDPoint();  
        origin2 = origin1;  
  
        // set the fields  
        origin1.x = 0.0;  
        origin1.y = 0.0;  
  
        // print  
        System.out.println(  
            "origin1 is at " + origin1.x + ", " + origin1.y);  
        System.out.println(  
            "origin2 is at " + origin2.x + ", " + origin2.y);  
    } // end main  
}  
} // end EqualPointPrinter
```

origin1 and origin2 are two different reference variables referring to the same point object.

3.8 Arrays

An array is a collection of variables of the same type. The args[] array of a main() method

is an array of Strings.

Consider a class which counts the occurrences of the digits 0-9. For example you might wish to test the randomness of a random number generator. If a random number generator is truly random, all digits should occur with equal frequency over a sufficiently long period of time.

You will do this by creating an array of ten ints called ndigit. The zeroth component of ndigit will track the number of zeros; the first component will track the numbers of ones and so forth. The RandomTest program below tests Java's random number generator to see if it produces apparently random numbers.

```
import java.util.Random;  
  
class RandomTest {  
  
    public static void main (String args[]) {  
  
        int[] ndigits = new int[10];  
  
        Random myRandom = new Random();  
  
        // Initialize the array  
        for (int i = 0; i < 10; i++) {  
            ndigits[i] = 0;  
        }  
  
        // Test the random number generator a whole lot  
        for (long i=0; i < 100000; i++) {  
            // generate a new random number between 0 and 9  
            double x = myRandom.nextDouble() * 10.0;  
            int n = (int) x;  
  
            //count the digits in the random number  
            ndigits[n]++;  
        }  
        // Print the results  
        for (int i = 0; i < 10; i++) {  
            System.out.println(i+": " + ndigits[i]);  
        }  
    }  
}
```

```
}
```

Below is one possible output from this program. If you run it your results should be slightly different. After all this is supposed to be random. These results are pretty much what you would expect from a reasonably random generator. If you have a fast CPU and some time to spare, try bringing the number of tests up to a billion or so, and see if the counts for the different digits get any closer to each other.

```
$ javac RandomTest.java
```

```
$ java RandomTest
```

```
0: 10171
```

```
1: 9724
```

```
2: 9966
```

```
3: 10065
```

```
4: 9989
```

```
5: 10132
```

```
6: 10001
```

```
7: 10158
```

```
8: 9887
```

```
9: 9907
```

There are three for loops in this program, one to initialize the array, one to perform the desired calculation, and a final one to print out the results. This is quite common in code that uses arrays.

3.9 Functions

Methods say what an object does.

```
class TwoDPoint {  
    double x;  
    double y;  
    void print() {  
        System.out.println(this.x + "," + this.y);  
    }  
}
```

Notice that you use the Java keyword `this` to reference a field from inside the same class.

```
TwoDPoint origin = new TwoDPoint();
origin.x = 0.0;
origin.y = 0.0;
origin.print();
```

noun-verb instead of verb-noun; that is subject-verb instead of verb-direct object.

subject-verb-direct object(s) is also possible.

3.10 Constructor

Constructors create new instances of a class, that is objects. Constructors are special methods that have the same name as their class and no return type. For example,

```
class TwoDPoint {
    double x;
    double y;
```

```
TwoDPoint(double xvalue, double yvalue) {
    this.x = xvalue;
    this.y = yvalue;
}
String getAsString() {
    return "(" + this.x + "," + this.y + ")";
}
void setX(double value) {
    this.x = value;
}
void setY(double value) {
    this.y = value;
}
double getX() {
    return this.x;
```

```
    }  
  
    double getY() {  
        return this.y;  
    }  
}
```

Constructors are used along with the new keyword to produce an object in the class (also called an instance of the class):

```
TwoDPoint origin = new TwoDPoint(0.0, 0.0);  
System.out.println("The x coordinate is " + origin.getX());
```

3.11 Errors in scripts

Web browsers are such an hostile environment that it is almost guaranteed that we will constantly deal with runtime errors. Users provide invalid input in ways you didn't think of. New browser versions change their behavior. An AJAX call fails for a number of reasons.

Many times we can't prevent runtime errors from happening, but at least we can deal with them in a manner that makes the user experience less traumatic.

Completely unhandled errors

Look at this seemingly trivial code snippet.

```
function getInput() {  
    var name = window.prompt('Type your name', '');  
    alert('Your name has ' + name.length + ' letters.');//  
}
```

It may not be obvious, but this code has a bug waiting to break free. If the user clicks Cancel or presses Esc the prompt() function will return null, which will cause the next line to fail with a null reference error.

If you as a programmer don't take any step to deal with this error, it will simply be delivered directly to the end user, in the form of a utterly useless browser error message like the one below.

Depending on the user's browser or settings, the error message may be suppressed and

only an inconspicuous icon shows up in the status bar. This can be worse than the error message, leaving the users thinking the application is unresponsive.

Globally handled errors

The window object has an event called onerror that is invoked whenever there's an unhandled error on the page.

```
window.onerror = function (message, url, lineNo) {  
    alert(  
        'Error: ' + message +  
        '\n Url: ' + url +  
        '\n Line Number: ' + lineNo);  
    return true;  
}
```

As you can see, the event will pass 3 arguments to the invoked function. The first one is the actual error message. The second one is the URL of the file containing the error (useful if the error is in an external .js file.) The last argument is the line number in that file where the error happened.

Returning true tells the browser that you have taken care of the problem. If you return false instead, the browser will proceed to treat the error as unhandled, showing the error message and the status bar icon.

Here's the message box that we will be showing to the user.

Structured Error Handling

The best way to deal with errors is to detect them the closest possible to where they happen. This will increase the chances that we know what to do with the error. To that effect JavaScript implements structured error handling, via the try...catch...finally block, also present in many other languages.

Syntax

```
try {  
    statements;
```

```
    } catch (error) {  
        statements;  
    } finally {  
        statements;  
    }
```

The idea is simple. If anything goes wrong in the statements that are inside the try block's statements then the statements in the catch block will be executed and the error will be passed in the error variable. The finally block is optional and, if present, is always executed last, regardless if there was an error caught or not.

Let's fix our example to catch that error.

```
function getInput(){  
    try {  
        var name = window.prompt('Type your name', "");  
        alert('Your name has ' + name.length + ' letters.');//  
    } catch (error) {  
        alert('The error was: ' + error.name +  
            '\n The error message was: ' + error.message);  
    } finally {  
        //do cleanup  
    }  
}
```

The error object has two important properties: name and message. The message property contains the same error message that we have seen before. The name property contains the kind of error that happened and we can use that to decide if we know what to do with that error.

It's a good programming practice to only handle the error on the spot if you are certain of what it is and if you actually have a way to take care of it (other than just suppressing it altogether.) To better target our error handling code, we will change it to only handle errors named "TypeError", which is the error name that we have identified for this bug.

```
function getInput(){
    try {
        var name = window.prompt('Type your name', '');
        alert('Your name has ' + name.length + ' letters.');
    } catch (error) {
        if (error.name === 'TypeError') {
            alert('Please try again.');
        } else {
            throw error;
        }
    } finally {
        //do cleanup
    }
}
```

Now if a different error happens, which is admittedly unlikely in this simple example, that error will not be handled. The throw statement will forward the error as if we never had this try...catch...finally block. It is said that the error will bubble up.

Throwing custom errors

We can use the throw statement to throw our own types of errors. The only recommendation is that our error object also has a name and message properties to be consistent in error handling.

```
throw {
    name: 'InvalidColorError',
    message: 'The given color is not a valid color value.'
};
```

Debugging

One of the most important activities in software development is debugging. It can also be one of the most costly. That's why we need to do our best to reduce the amount of time spent in debugging.

One way to reduce this time is to create automated unit tests, which we will see in the lesson Production Grade JavaScript.

Another way is to use the best tools available and try to remove the pain associated with debugging. It used to be the case that debugging tools for JavaScript were archaic or close to non-existent. This situation has improved a lot and now we can confidently say we have feasible ways to debug JavaScript without resorting to horrendous tactics, such as sprinkling alert() calls across our code.

We won't waste your time discussing all the existing tools for debugging. Instead we will focus on the tool that singlehandedly re-wrote the JavaScript debugging history.

UNIT - 4

JAVASCRIPT AND HTML DOCUMENTS

4.1 The Javascript execution environment

Features

The following features are common to all conforming ECMAScript implementations, unless explicitly specified otherwise.

Imperative and structured

JavaScript supports all the structured programming syntax in C (e.g., if statements, while loops, switch statements, etc.). One partial exception is scoping: C-style block-level scoping is not supported (instead, JavaScript has function-level scoping). JavaScript 1.7, however, supports block-level scoping with the let keyword. Like C, JavaScript makes a distinction between expressions and statements. One syntactic difference from C is automatic semicolon insertion, in which the semicolons that terminate statements can be omitted.^[18]

Dynamic

dynamic typing

As in most scripting languages, types are associated with values, not variables.

For example, a variable x could be bound to a number, then later rebound to a string. JavaScript supports various ways to test the type of an object, including duck typing.^[19]

object based

JavaScript is almost entirely object-based. JavaScript objects are associative arrays, augmented with prototypes (see below). Object property names are string keys: obj.x = 10 and obj["x"] = 10 are equivalent, the dot notation being syntactic sugar. Properties and their values can be added, changed, or deleted at run-time. Most properties of an object (and those on its prototype inheritance chain) can be enumerated using a for...in loop. JavaScript has a small number of built-in objects such as Function and Date.

run-time evaluation

JavaScript includes an eval function that can execute statements provided as strings at run-time.

Functional

first-class functions

Functions are first-class; they are objects themselves. As such, they have properties and can be passed around and interacted with like any other object.

inner functions and closures

Inner functions (functions defined within other functions) are created each time the outer function is invoked, and variables of the outer functions for that invocation continue to exist as long as the inner functions still exist, even after that invocation is finished (e.g. if the inner function was returned, it still has access to the outer function's variables) — this is the mechanism behind closures within JavaScript.

Prototype-based

prototypes

JavaScript uses prototypes instead of classes for inheritance. It is possible to simulate many class-based features with prototypes in JavaScript.

functions as object constructors

Functions double as object constructors along with their typical role. Prefixing a function call with new creates a new object and calls that function with its local this keyword bound to that object for that invocation. The constructor's prototype property determines the object used for the new object's internal prototype. JavaScript's built-in constructors, such as Array, also have prototypes that can be modified.

functions as methods

Unlike many object-oriented languages, there is no distinction between a function definition and a method definition. Rather, the distinction occurs during function calling; a function can be called as a method. When a function is called as a method of an object, the function's local this keyword is bound to that object for that invocation.

Miscellaneous

run-time environment

JavaScript typically relies on a run-time environment (e.g. in a web browser) to provide objects and methods by which scripts can interact with "the outside world". In fact, it relies on the environment to provide the ability to include/import scripts (e.g. HTML <script> elements). (This is not a language feature per se, but it is common in most JavaScript implementations.)

variadic functions

An indefinite number of parameters can be passed to a function. The function can access them through formal parameters and also through the local arguments object.

array and object literals

Like many scripting languages, arrays and objects (associative arrays in other languages) can each be created with a succinct shortcut syntax. In fact, these literals form the basis of the JSON data format.

regular expressions

JavaScript also supports regular expressions in a manner similar to Perl, which provide a concise and powerful syntax for text manipulation that is more sophisticated than the built-in string functions.

4.2 The Document Object Model

The **Document Object Model (DOM)** is a cross-platform and language-independent convention for representing and interacting with objects in HTML, XHTML and XML documents. Aspects of the DOM (such as its "Elements") may be addressed and manipulated within the syntax of the programming language in use. The public interface of a DOM is specified in its Application Programming Interface (API).

Applications

DOM is likely to be best suited for applications where the document must be accessed repeatedly or out of sequence order. If the application is strictly sequential and one-pass, the SAX model is likely to be faster and use less memory. In addition, non-extractive XML parsing models such as VTD-XML, provide a new memory-efficient option.

Web browsers

A web browser is not obliged to use DOM in order to render an HTML document. However, the DOM is required by JavaScript scripts that wish to inspect or modify a web page dynamically. In other words, the Document Object Model is the way JavaScript sees its containing HTML page and browser state.

Implementations

Because DOM supports navigation in any direction (e.g., parent and previous sibling) and allows for arbitrary modifications, an implementation must at least buffer the document that has been read so far (or some parsed form of it).

Layout engines

Web browsers rely on layout engines to parse HTML into a DOM. Some layout engines such as Trident/MSHTML and Presto are associated primarily or exclusively with a particular browser such as Internet Explorer and Opera respectively. Others, such as WebKit and Gecko, are shared by a number of browsers, such as Safari, Google Chrome, Firefox or Flock. The different layout engines implement the DOM standards to varying degrees of compliance.

See also: Comparison of layout engines (Document Object Model)

Libraries

DOM implementations:

- libxml2
- MSXML
- Xerces is a C++ DOM implementation with Java and Perl bindings

APIs that expose DOM implementations:

- JAXP (Java API for XML Processing) is an API for accessing DOM providers

Alternative non-DOM tree-based XML libraries:

- VTD-XML is a Java library that offers alternative tree-based view to XML documents

4.3 Element access in Javascript

Javascript provides the ability for getting the value of an element on a webpage as well as dynamically changing the content within an element.

Getting the value of an element

To get the value of an element, the **getElementById** method of the **document** object is used. For this method to get the value of an element, that element has to have an id given to it through the **id** attribute.

Example:

```
<script type="text/javascript"> function getText(){ //access the element with the id 'textOne' and get its value //assign this value to the variable theText var theText = document.getElementById('textOne').value; alert("The text in the textbox is " + theText); } </script> <input type="text" id="textOne" /> <input type="button" value="Get text" onclick="getText()" />
```

Changing the content within an element

To change the content within an element, use the **innerHTML** property. Using this property, you could replace the text in paragraphs, headings and other elements based on several things such as a value the user enters in a textbox. For this property to change the content within an element, that element has to have an 'id' given to it through the **id** attribute.

Example:

```
<script type="text/javascript"> function changeTheText(){ //change the innerHTML property of the element whose id is 'text' to 'So is HTML!' document.getElementById('text').innerHTML = 'So is HTML!'; } </script> <p id="text">Javascript is cool!</p> <input type='button' onclick='changeTheText()' value='Change the text' />
```

You can also change the text of elements based on user input:

Example:

```
<script type="text/javascript"> function changeInputText(){ /* change the innerHTML property of the element whose id is 'theText' to the value from the variable usersText which will take the value from the element whose id is 'usersText' */ var usersText = document.getElementById('usersText').value; document.getElementById('theText').innerHTML = usersText; } </script> <p id="theText">Enter some text into the textbox and click the button</p> <input type="text" id="usersText" /> <br /> <input type="button" onclick="changeInputText()"
```

value="Change the text" />

4.4 Events and event handling

HTML Events

Not all events are significant to a program. We can simplify situations to determine which events are significant, and which events we can leave out. For example, a leaf hitting the bus is an event, but is not a significant event if we wish to model the cost effectiveness of the bus route.

As we know, HTML provides elements to allow us to create web documents. We can fill these elements with our own data and display unique web documents. We can position the elements with careful planning, and create aesthetic web sites.

This is the object-orientated part of web documents. DHTML, or Dynamic HTML supplies the event-driven side of things.

To begin using DHTML, and therefore event-driven HTML, we need to look at what events happen to our web document.

There are over 50 events in all. Some can happen to a lot of HTML elements, some only happen to specific HTML elements.

For example, moving the mouse pointer and clicking the mouse buttons create the following events:

- onmousedown
- onmousemove
- onmouseout
- onmouseover
- onmouseup
- onclick
- ondblclick

The event happens to the object. So, if the mouse pointer is over an image, say Image1, and we click it, we create an (object, event) pair, in this case (Image1, onclick). If we move the mouse into a select box, Select1, we create (Select1, onmouseover), and so-on.

We need technical definitions of the events to be able to appreciate what is actually happening.

Mouse Event	Description
onmousedown	A mouse button has been pressed
onmousemove	The mouse has been moved
onmouseout	The mouse pointer has left an element
onmouseover	The mouse pointer has entered an element
onmouseup	A mouse button has been released
onclick	A mouse button has been clicked
ondblclick	A mouse button has been double-clicked (clicked twice rapidly)

The idea behind this is to represent all the possibilities we may encounter in computer events. This doesn't allow for events like the mouse is dirty, or the mouse is upside-down

- these are not relevant.

You may also notice that some actions will fire two events. Moving the mouse into or out of an element will fire both a mousemove event and a mouseover (or mouseout) event. Clicking a mouse button fires a mousedown, click and mouseup events.

The other user-type event is generated by the keyboard.

Keyboard Event	Description
onkeydown	A key has been pressed
onkeypress	onkeydown followed by onkeyup
onkeyup	A key has been released

Remaining events have a more conceptual taste. There are a few too many to list here, but they are well documented. Commonly used events include:

Event	Description
onblur	An element loses focus

onerror	An error occurs
onfocus	An element gains focus
onload	The document has completely loaded
onreset	A form reset command is issued
onscroll	The document is scrolled
onselect	The selection of element has changed
onsubmit	A form submit command is issued

We will look at the onfocus event at the end of this article, which may help your understanding of the subtler events.

Attaching Events to HTML elements

The first thing you need to know is that when onblur or onmouseover or any other event is used within your HTML page, it is commonly referred to as an event handler.

Now we need to write some JavaScript to enhance the current functionality of the element, and this is where it can get complicated.

We all know what a typical HTML element looks like, and hopefully we know what an HTML element with the STYLE attribute looks

like. As a re-cap, an HTML element is:

```
<P>Hello from the 60's</P>
```

and with style:

```
<P STYLE="position:absolute;top:10;left:10;color:black">Hello from the 60's</P>
```

To attach the event handler of our choice to this <P> element, we need to notify the element of the type of event to watch out for, and also we have to tell the element what to do when it does receive the event.

To create the psychedelic 60's effect I have in mind, we need to monitor for the onmousemove event. The JavaScript to create this code is fairly simple, we change the color to a random one every time the mouse pointer moves over the <P> element.

The whole code for the <P> element therefore looks like:

```
<P onmousemove="style.color=Math.floor(Math.random()*16777215);"  
STYLE="position:absolute;top:10;left:10;color:black">Hello from the 60's</P>
```

Placed in an HTML document, we get:

```
<HTML>  
<HEAD>  
<TITLE>Hello from the 60's</TITLE></HEAD>  
<BODY>  
<p onmousemove="style.color=Math.floor(Math.random()*16777216);"  
STYLE="position:absolute;top:10;left:10;color:black">Hello from the 60's</P>  
</BODY>  
</HTML>
```

When we move the mouse pointer over the text, we get a very psychedelic effect. Note that the mouse pointer must be moved, it's mere presence over the text is not an event.

The strange number, 16777216, is the total number of colors we get. We can use any of 256 red shades, 256 green shades and 256 blue shades, so the total is $256^3 = 16777216$.

This is slightly obscured by the fact that the first number is 0, and so we only want the range (0 - 16777215). This is created by the Math.floor() method with the Math.random() method. Math.random() returns a number between greater than or equal to zero, but less than 1. When we floor the random function, we will never get 16777216, but will get 0 - 16777215.

Other Event Handling Techniques

This is not the only method we have for attaching events to elements, and nor is it the

only method we have for implementing JavaScript.

Function Calls

We could house the JavaScript in a function, and so the 60's program becomes:

```
<HTML>
<HEAD>
<TITLE>Hello from the 60's Again</TITLE>
</HEAD>
<SCRIPT>
function randomcolor(){
event.srcElement.style.color=Math.floor(Math.random()*16777216);
}
</SCRIPT>
<BODY>
<Ponmousemove="randomcolor();"
STYLE="position:absolute;top:10;left:10;color:black">Hello from the 60's</P>
<P
onmousemove="randomcolor();STYLE="position:absolute;top:50;left:20;color:black">
Hello from the 70's</P></BODY></HTML>
```

Now our event handler calls the randomcolor() function.

The randomcolor() function is slightly more complicated than the method used before. It makes use of the Event object, which I will cover in detail in a later article.

One of the properties of the event object is the srcElement property, which contains the element that initiated the event.

We can use this style of event handler to make our code terser and more efficient. We can even use the same function twice, and independently.

Attaching Events

We also have several different techniques for attaching the event to an object. We have met one of them already, which is direct attachment of the event to an object.

We can also indirectly attach the event to the object. We can achieve the same event handling effect as before from within a script tag.

```
<HTML>
<HEAD>
<TITLE>Hello from the 60's Part 3</TITLE>
</HEAD>
<BODY>
<P ID="sixties" STYLE="position:absolute;top:10;left:10;color:black">Hello from the
60's</P>
</BODY>
</HTML>
<SCRIPT>
Function sixties.onmousemove() {
event.srcElement.style.color=Math.floor(Math.random()*16777216);
}
</SCRIPT>
```

This code is fairly stylish, but causes a few more problems. As the script contains a reference to the sixties object, this object must have been created before the script can use it. If we place the **<SCRIPT>** element where we usually do, inside the **<HEAD>** element, we get an error message, because the sixties object does not yet exist. So we must place it after the sixties object has been created.

We can also create little script elements, which we assign to an object.

```
<HTML>
<HEAD>
<TITLE>Hello from the 60's IV</TITLE>
</HEAD>
<BODY>
<P ID="sixties" STYLE="position:absolute;top:10;left:10;color:black">Hello from the
60's</P>
</BODY>
</HTML>
<SCRIPT FOR="sixties" EVENT="onmousemove">
event.srcElement.style.color=Math.floor(Math.random()*16777216);
</SCRIPT>
```

To finish the article, we shall look at one of the more conceptual events, the focus event.

4.6 Button elements

Syntax

<BUTTON>...</BUTTON>

Attribute

Specifications

- NAME=CDATA (key in submitted form)
- VALUE=CDATA (value in submitted form)
- TYPE=[submit | reset | button] (type of button)
- DISABLED (disable button)
 - ACCESSKEY=Character (shortcut key)
 - TABINDEX=Number (position in tabbing order)
 - ONFOCUS=Script (element received focus)
 - ONBLUR=Script (element lost focus)
 - common attributes

Contents

- Inline elements except A, INPUT, SELECT, TEXTAREA, LABEL, BUTTON, and IFRA
ME
- Block-level elements except FORM, ISINDEX, and FIELDSET

Contained in

Block-level elements, inline elements except BUTTON

The **BUTTON** element defines a submit button, reset button, or push button. Authors can also use **INPUT** to specify these buttons, but the **BUTTON** element allows richer labels, including images and emphasis. However, **BUTTON** is new in HTML 4.0 and not as widely supported as **INPUT**. For compatibility with old browsers, **INPUT** should generally be used instead of **BUTTON**.

The **TYPE** attribute of **BUTTON** specifies the kind of button and takes the value **submit** (the default), **reset**, or **button**. The **NAME** and **VALUE** attributes determine the name/value pair sent to the server when a submit button is pushed. These attributes allow authors to provide multiple submit buttons and have the form handler take a different action depending on the submit button used.

Some examples of **BUTTON** follow:

- <BUTTON NAME=submit VALUE=modify ACCESSKEY=M>Modify

```
information</BUTTON>
<BUTTON NAME=submit VALUE=continue ACCESSKEY=C>Continue with
application</BUTTON>
• <BUTTON ACCESSKEY=S>Submit <IMG SRC="checkmark.gif"
ALT=""></BUTTON>
<BUTTON TYPE=reset ACCESSKEY=R>Reset <IMG SRC="x.gif"
ALT=""></BUTTON>
• <BUTTON TYPE=button ID=toggler ONCLICK="toggle()"
ACCESSKEY=H>Hide <strong>non-strict</strong> attributes</BUTTON>
```

The **ACCESSKEY** attribute, used throughout the preceding examples, specifies a single Unicode character as a shortcut key for pressing the button. Entities (e.g. **´**) may be used as the **ACCESSKEY** value.

The boolean **DISABLED** attribute makes the **BUTTON** element unavailable. The user is unable to push the button, the button cannot receive focus, and the button is skipped when navigating the document by tabbing.

The **TABINDEX** attribute specifies a number between 0 and 32767 to indicate the tabbing order of the button. A **BUTTON** element with **TABINDEX=0** or no **TABINDEX** attribute will be visited after any elements with a positive **TABINDEX**. Among positive **TABINDEX** values, the lower number receives focus first. In the case of a tie, the element appearing first in the HTML document takes precedence.

The **BUTTON** element also takes a number of attributes to specify client-side scripting actions for various events. In addition to the core events common to most elements, **BUTTON** accepts the following event attributes:

- **ONFOCUS**, when the element receives focus;
- **ONBLUR**, when the element loses focus.

4.7 Text box and Password elements

Introduction to forms

In HTML form is a section of a document containing normal content, markup, special elements called controls (checkboxes, radio buttons, menus, etc.), and labels on those controls. Users generally "complete" a form by modifying its controls (entering text, selecting menu items, etc.), before submitting the form to an agent for processing (e.g., to a Web server, to a mail server, etc.)

Here's a simple form that includes labels, radio buttons, and push buttons (reset the form or submit it):

```
<FORM action="http://somesite.com/prog/adduser" method="post">  
  <P>  
    <LABEL for="firstname">First name: </LABEL>  
      <INPUT type="text" id="firstname"><BR>  
    <LABEL for="lastname">Last name: </LABEL>  
      <INPUT type="text" id="lastname"><BR>  
    <LABEL for="email">email: </LABEL>  
      <INPUT type="text" id="email"><BR>  
    <INPUT type="radio" name="sex" value="Male"> Male<BR>  
    <INPUT type="radio" name="sex" value="Female"> Female<BR>  
    <INPUT type="submit" value="Send"> <INPUT type="reset">  
  </P>  
</FORM>
```

4.8 The DOM 2 event model

The DOM2 Event model specification (<http://www.w3.org/TR/DOM-Level-2-Events/>) describes a standard way to create, capture, handle, and cancel events in a tree-like structure such as an (X)HTML document's object hierarchy. It also describes event propagation behavior, that is, how an event arrives at its target and what happens to it afterward.

The DOM2 approach to events accommodates the basic event model and marries important concepts from the proprietary models. This essentially means that the basic event model works exactly as advertCSEd in a DOM2-supporting browser.

Also, everything that you can do in Netscape 4 and Internet Explorer you can do in a DOM2 browser, but the syntax is different.

The hybridization of the proprietary models is evident in how events propagate in a DOM2-supporting browser. Events begin their lifecycle at the top of the hierarchy (at the **Document**) and make their way down through containing objects to the target. This is known as the capture phase because it mimics the behavior of Netscape 4. During its descent, an event may be pre-processed, handled, or redirected by any intervening object. Once the event reaches its target and the handler there has executed, the event proceeds back up the hierarchy to the top. This is known as the bubbling phase because of its obvious connections to the model of Internet Explorer 4+.

Mozilla-based browsers were the first major browsers to implement the DOM2 Events standard. These browsers include Mozilla itself, Netscape 6+, Firefox, Camino, and others. Opera 7 has nearly complete support, as does Safari (a popular MacOS browser). In fact, most browsers (with the exception of those from Microsoft) support or will soon support DOM2 Events. This is as it should be; uniformity of interface is the reason we have standards in the first place.

The fly in the ointment is that, as of version 6, Internet Explorer doesn't support DOM2 Events. Microsoft does not appear to have plans to add support in the near future, and it's unclear whether they plan to in the medium- or long-term. So, unfortunately, Web programmers aren't likely to be free of the headache of cross-browser scripting for events any time soon and should be wary of focusing solely on the DOM2 Event specification.

Binding Handlers to Objects

The easiest way to bind event handlers to elements under DOM Level 2 is to use the (X)HTML attributes like **onclick** that you should be familiar with by now. Nothing changes for DOM2-supporting browsers when you bind events in this way, except that only support for events in the (X)HTML standard is guaranteed (though some browsers support more events).

Because there is no official DOM2 way for script in the text of event handler attributes to access an **Event** object, the preferred binding technique is to use

JavaScript. The same syntax is used as with the basic event model:

```
<<p id="myElement">>Click on me<</p>>
<<p>>Not on me<</p>>
<<script type="text/javascript">>
<<!--
function handleClick(e)
{
    alert("Got a click: " + e);
    // IE5&6 will show an undefined in alert since they are not DOM2
}
document.getElementById("myElement").onclick = handleClick;
//-->>
<</script>>
```

Notice in this example how the handler accepts an argument. DOM2 browsers pass an **Event** object containing extra information about the event to handlers. The name of the argument is arbitrary, but `-event`, `-e`, and `-evnt` are most commonly used. We'll discuss the **Event** object in more detail in an upcoming section.

DOM2 Event Binding Methods

You can also use the new **addEventListener()** method introduced by DOM2 to engage an event handler in a page. There are three reasons you might wish to use this function instead of directly setting an object's event handler property. The first is that it enables you to bind multiple handlers to an object for the same event. When handlers are bound in this fashion, each handler is invoked when the specified event occurs, though the order in which they are invoked is arbitrary. The second reason to use **addEventListener()** is that it enables you to handle events during the capture phase (when an event `trickles down` to its target). Event handlers bound to event handler attributes like **onclick** and **onsubmit** are only invoked during the bubbling phase. The third reason is that this method enables you to bind handlers to text nodes, an impossible task prior to DOM2.

The syntax of the **addEventListener()** method is as follows:

object.addEventListener(`-event-`, `handler`, `capturePhase`);

- `object` is the node to which the listener is to be bound.

- "event" is a string indicating the event it is to listen for.
- handler is the function that should be invoked when the event occurs.
- capturePhase is a Boolean indicating whether the handler should be invoked during the capture phase (**true**) or bubbling phase (**false**).

For example, to register a function `changeColor()` as the capture-phase **mouseover** handler for a paragraph with **id** of `myText` you might write

```
document.getElementById('myText').addEventListener("mouseover",
changeColor, true);
```

To add a bubble phase handler `swapImage()`:

```
document.getElementById('myText').addEventListener("mouseover",
swapImage, false);
```

Handlers are removed using **removeEventListener()** with the same arguments as given when the event was added. So to remove the first handler in the previous example (but keep the second) you would invoke

```
document.getElementById('myText').removeEventListener("mouseover",
changeColor, true);
```

We'll see some specific examples using the **addEventListener()** later on in the chapter.

Event Objects

As previously mentioned, browsers supporting DOM2 Events pass an **Event** object as an argument to handlers. This object contains extra information about the event that occurred, and is in fact quite similar to the **Event** objects of the proprietary models. The exact properties of this object depend on the event that occurred, but all **Event** objects have the read-only properties listed in Table 11-10.

Table 11-10: Properties Common to All Event Objects

Read-Only Property	Description
<code>>bubbles</code>	Boolean indicating whether the event bubbles

>cancelable	Boolean indicating whether the event can be canceled
>currentTarget	Node whose handler is currently executing (i.e., the node the handler is bound to)
>eventPhase	Numeric value indicating the current phase of the event flow (1 for capture, 2 if at the target, 3 for bubble)
>type	String indicating the type of the event (such as "click")
>target	Node to which the event was originally dispatched (i.e., the node at which the event occurred)

We list the properties specific to each event in the following sections as we discuss the different kinds of events DOM2-supporting browsers enable you to handle.

Mouse Events

The mouse events defined by DOM2 are those from (X)HTML. They're listed in Table 11-11. Since, under DOM2, not all events include a bubbling phase and all default actions can be canceled, Table also lists these behaviors for each event.

Table : Mouse-Related Events Supported Under DOM2 Events

Event	Bubbles?	Cancelable?
click	Yes	Yes
mousedown	Yes	Yes
mouseup	Yes	Yes
mouseover	Yes	Yes

mousemove	Yes	No
mouseout	Yes	Yes

When a mouse event occurs, the browser fills the **Event** object with the extra information shown in Table

Table : Additional Properties of the Event Object When the Event Is Mouse-Related

Property	Description
>altKey	Boolean indicating if the ALT key was depressed
>button	Numeric value indicating which mouse button was used (typically 0 for left,

Table : Additional Properties of the Event Object When the Event Is Mouse-Related

Property	Description
	1 for middle, 2 for right)
>clientX	Horizontal coordinate of the event relative to the browser's content pane
>clientY	Vertical coordinate of the event relative to the browser's content pane
>ctrlKey	Boolean indicating if the CTRL key was depressed during event
>detail	Indicating the number of times the mouse button was clicked (if at all)
>metaKey	Boolean indicating if the META key was depressed during event

>relatedTarget	Reference to a node related to the event—for example, on a mouseover it references the node the mouse is leaving; on mouseout it references the node to which the mouse is moving
>screenX	Horizontal coordinate of the event relative to the whole screen
>screenY	Vertical coordinate of the event relative to the whole screen
>shiftKey	Boolean indicating if the SHIFT key was depressed during event

The following example illustrates their use:

```
<<!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>>
<<title>>DOM2 Mouse Events<</title>>
<<meta http-equiv="content-type" content="text/html; charset=ISO-8859-1" />>
<</head>>
<<body>>
<<h2>>DOM2 Mouse Events<</h2>>
<<form id="mouseform" name="mouseform" action="#" method="get">>
Alt Key down?
<<input id="altkey" type="text" />><<br />>
Control Key down?
<<input id="controlkey" type="text" />><<br />>
Meta Key down?
<<input id="metakey" type="text" />><<br />>
Shift Key down?
```

```
<<input id="shiftkey" type="text" />><<br />>
```

Browser coordinates of click: <<input id="clientx" type="text" />>,
<<input id="clienty" type="text" />><
>

Screen coordinates of click: <<input id="screenx" type="text" />>,
<<input id="screeny" type="text" />><
>

Button used: <<input id="buttonused" type="text" />><
><
>
<</form>>
<<hr />>

Click anywhere on the document...

```
<<script type="text/javascript">>
```

```
<<!--
```

```
function showMouseDetails(event)
```

```
{
```

```
    var theForm = document.mouseform;  
    theForm.altkey.value = event.altKey;  
    theForm.controlkey.value = event.ctrlKey;  
    theForm.shiftkey.value = event.shiftKey;  
    theForm.metakey.value = event.metaKey;  
    theForm.clientx.value = event.clientX;  
    theForm.clienty.value = event.clientY;  
    theForm.screenx.value = event.screenX;  
    theForm.screeny.value = event.screenY;  
    if (event.button == 0)  
        theForm.buttonused.value = "left";  
    else if (event.button == 1)  
        theForm.buttonused.value = "middle";  
    else  
        theForm.buttonused.value = "right";
```

```
}
```

```
document.addEventListener("click", showMouseDetails, true);
```

```
//-->>
```

```
<</script>>
<</body>>
<</html>>
```

The result of a click is shown in Figure .

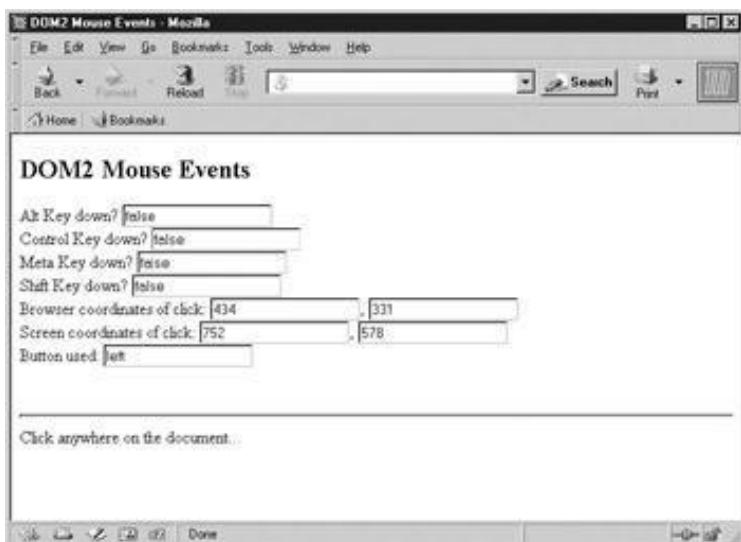


Figure: Contextual information is passed in through the **Event** object.

Keyboard Events

Surprisingly, DOM Level 2 does not define keyboard events. They will be specified in a future standard, the genesis of which you can see in DOM Level 3. Fortunately, because (X)HTML allows **keyup**, **keydown**, and **keypress** events for many elements, you'll find that browsers support them. Furthermore, with IE dominating the picture, you still have that event model to fall back on. Table lists the keyboard-related events for DOM2-compliant browsers, as well as their behaviors.

Table: Keyboard Events Supported by Most Browsers

Event	Bubbles?	Cancelable?
keyup	Yes	Yes
keydown	Yes	Yes
keypress	Yes	Yes

The Mozilla-specific key-related properties of the **Event** object are listed in Table .

Table: Additional Properties of the Event Object for Key-Related Events in Mozilla

Property	Description
>altKey	Boolean indicating if the ALT key was depressed
>charCode	For printable characters, a numeric value indicating the Unicode value of the key depressed
>ctrlKey	Boolean indicating if the CTRL key was depressed during event
>isChar	Boolean indicating whether the keypress produced a character (useful because some key sequences such as CTRL-ALT do not)
>keyCode	For non-printable characters, a numeric value indicating the Unicode value of the key depressed
>metaKey	Boolean indicating if the META key was depressed during event
>shiftKey	Boolean indicating if the SHIFT key was depressed during event

Browser Events

DOM2 browsers support the familiar browser and form-related events found in all major browsers. The list of these events is found in Table .

Table: Browser- and Form-Related DOM2 Events and Their Behaviors

Event	Bubbles?	Cancelable?
load	No	No

Table: Browser- and Form-Related DOM2 Events and Their Behaviors

Event	Bubbles?	Cancelable?
unload	No	No
abort	Yes	No
error	Yes	No
select	Yes	No
change	Yes	No
submit	Yes	Yes
reset	Yes	No
focus	No	No
Blur	No	No
resize	Yes	No
scroll	Yes	No

UI Events

Although DOM Level 2 builds primarily on those events found in the (X)HTML specification (and DOM Level 0), it adds a few new User Interface (UI) events to round out the field. These events are prefixed with `-DOM-` to distinguish them from `-normal-` events. These events are listed in Table.

Table: UI-Related DOM2 Events and Their Behaviors

Event	Bubbles?	Cancelable?
DOMFocusIn	Yes	No
DOMFocusOut	Yes	No
DOMActivate	Yes	Yes

The need for and meaning of these events is not necessarily obvious. **DOMFocusIn** and **DOMFocusOut** are very similar to the traditional **focus** and **blur** events, but can be applied to any element, not just form fields. The **DOMActivate** event is fired when an object is receiving activity from the user. For example, it fires on a link when it is clicked and on a select menu when the user activates the pull-down menu. This event is useful when you don't care how the user invokes the element's functionality, just that it is being used. For example, instead of using both an **onclick** and **onkeypress** handler to trap link activation (via the mouse or keyboard) you could register to receive the **DOMActivate** event. While these new events are rarely used, it is helpful to be aware of them should you encounter them in new scripts.

Event Creation

The last DOM 2 Event topic we mention is not often used nor implemented in browsers, but is interesting nonetheless—event creation. The DOM2 Event specification allows for synthetic events to be created by the user using **document.createEvent()**. You first create the type of event you want, say an HTML-related event:

```
evt = document.createEvent("HTMLEvents");
```

Then once your event is created you pass it various attributes related to the event type. Here, for example, we pass the type of event "click" and Boolean values indicating it is bubble-able and cancelable:

```
evt.initEvent("click", "true", "true");
```

Finally, we find a node in the document tree and dispatch the event to it:

```
currentNode.dispatchEvent(evt);
```

The event then is triggered and reacts as any other event.

The following example shows DOM2 event creation in action and allows you to move around the tree and fire clicks at various locations. **The addEventListener()** and **removeEventListener()** are added into the example so you do not have to observe click events until you are ready.

```
<<!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>>
<<title>>DOM2 Event Creation<</title>>
<<meta http-equiv="content-type" content="text/html; charset=ISO-8859-1" />>
<</head>>
<<body>>
<<h2>>DOM2 Event Creation<</h2>>
<<form id="mouseform" name="mouseform" action="#" method="get">>
Browser coordinates of click: <<input id="clientx" type="text" />>,
<<input id="clienty" type="text" />> <<br />>
<</form>>
<<br />><<hr />><<br />>
<<script type="text/javascript">>
<<!--
// DOM 2 Only - no IE6 support
```

```
function showMouseDetails(event)
{
    document.mouseform.clientx.value = event.clientX;
    document.mouseform.clienty.value = event.clientY;
}
function makeEvent()
{
    evt = document.createEvent("HTMLEvents");
    evt.initEvent("click","true","true");
    currentNode.dispatchEvent(evt);
}
function startListen()
{
```

```
        document.addEventListener("click", showMouseDetails, true);  
    }  
  
    function stopListen()  
    {  
        document.removeEventListener("click", showMouseDetails, true);  
    }  
  
    startListen();  
//-->>  
<</script>>  
<<form action="#" method="get" id="myForm" name="myForm">>  
    Current Node: <<input type="text" name="statusField" value="" />>  
    <<br />>  
    <<input type="button" value="parent" onclick="if  
(currentNode.parentNode) currentNode = currentNode.parentNode;  
document.myForm.statusField.value = currentNode.nodeName;" />>  
    <<input type="button" value="First Child" onclick="if  
(currentNode.firstChild) currentNode = currentNode.firstChild;  
document.myForm.statusField.value = currentNode.nodeName;" />>  
    <<input type="button" value="Next Sibling" onclick="if  
(currentNode.nextSibling) currentNode = currentNode.nextSibling;  
document.myForm.statusField.value = currentNode.nodeName;" />>  
    <<input type="button" value="Previous Sibling" onclick="if  
(currentNode.previousSibling) currentNode = currentNode.previousSibling;  
document.myForm.statusField.value = currentNode.nodeName;" />>  
    <<br />><<br />>  
    <<input type="button" value="Start Event Listener"  
onclick="startListen();" />>  
    <<input type="button" value="Stop Event Listener"  
onclick="stopListen();" />>  
    <<br />><<br />>  
    <<input type="button" value="Create Event" onclick="makeEvent();" />>  
<</form>>
```

```
<<script type="text/javascript">>  
<<!--  
var currentNode = document.body;  
document.myForm.statusField.value = currentNode.nodeName;  
//-->>  
<</script>>  
<</body>>  
<</html>>
```

There are a number of details to DOM2 Event creation that we forgo primarily because it is not widely implemented in browsers. In fact with much of this section it should always be kept in mind that the DOM2 currently is probably not the best approach to making events work across browsers since it is not supported by Internet Explorer. We review the sorry state of affairs for event support in browsers briefly now.

4.9 The navigator object

The JavaScript navigator object is the object representation of the client internet browser or web navigator program that is being used. This object is the top level object to all others.

Navigator Properties

- appCodeName - The name of the browser's code such as "Mozilla".
- appMinorVersion - The minor version number of the browser.
- appName - The name of the browser such as "Microsoft Internet Explorer" or "Netscape Navigator".
- appVersion - The version of the browser which may include a compatibility value and operating system name.
- cookieEnabled - A boolean value of true or false depending on whether cookies are enabled in the browser.
- cpuClass - The type of CPU which may be "x86"
- mimeTypes - An array of MIME type descriptive strings that are supported by the

browser.

- onLine - A boolean value of true or false.
- opsProfile
- platform - A description of the operating system platform. In my case it is "Win32" for Windows 95.
- plugins - An array of plug-ins supported by the browser and installed on the browser.
- systemLanguage - The language being used such as "en-us".
- userAgent - In my case it is "Mozilla/4.0 (compatible; MSIE 4.01; Windows 95)" which describes the browser associated user agent header.
- userLanguage - The language the user is using such as "en-us".
- userProfile

Methods

- javaEnabled() - Returns a boolean telling if the browser has JavaScript enabled.
- taintEnabled() - Returns a boolean telling if the browser has tainting enabled.
Tainting is a security protection mechanism for data.

Navigator Objects

Most Navigator objects have corresponding HTML tags. Refer to the appropriate object section for more information.

- MimeType - Allows access to information about MIME types supported by the browser.
- plugin - Access to information about all plugins the browser supports. <EMBED>
- window - The browser frame or window.
 - frame- Allows access to all frames in the window. <FRAME>
 - history - The URLs previously accessed in a window.
 - location - Represents a URL.
 - document - The HTML document loaded in the window. <BODY>
 - anchor object anchors array - Allows access to all anchors in the document.
 - applet - Allows access to all applets in the document.
 - area - Allows access to an area in a client image map. <MAP>
 - image object images array - Allows access to all images in the

document.

- link object links array - Allows access to all links in the document.

- layer - Allows access to HTML layers.
- forms - Allows access to all forms in the document. <FORM>
 - button - Allows access to a form button exclusive of a submit or reset button. <INPUT TYPE="button">
 - checkbox - Allows access to a formcheckbox. <INPUT TYPE="checkbox">
 - element - Allows access to buttons or fields in a form.
 - FileUpLoad - Allows access to a form file upload element.
<INPUT TYPE="file">
 - hidden - Allows access to a form hidden field. <INPUT TYPE="hidden">
 - option
 - password - Allows access to a form password field.
<INPUT TYPE="password">
 - radio - Allows access to a form radio button set. <INPUT TYPE="radio">
 - reset - Allows access to a form reset button.. <INPUT TYPE="reset">
 - select - Allows access to a form selected list with the option allowing access to selected elements in the select list.
<SELECT>
 - submit - Allows access to a form submit button. <INPUT TYPE="submit">
 - text - Allows access to a form text field. <INPUT TYPE="text">
 - textarea - Allows access to a form text area field.
<TEXTAREA>
- embeds - Allows access to embedded plug ins.

4.10 DOM tree traversal and modification.

Its TreeWalker, NodeIterator, and NodeFilter interfaces provide easy-to-use, robust, selective traversal of a document's contents.

The interfaces found within this section are not mandatory. A DOM application may use the hasFeature(feature, version) method of the DOMImplementation interface with parameter values "Traversal" and "2.0" (respectively) to determine whether or not this module is supported by the implementation. In order to fully support this module, an implementation must also support the "Core" feature defined in the DOM Level 2 Core specification [DOM Level 2 Core]. Please refer to additional information about conformance in the DOM Level 2 Core specification [DOM Level 2 Core].

NodeIterators and TreeWalkers are two different ways of representing the nodes of a document subtree and a position within the nodes they present. A NodeIterator presents a flattened view of the subtree as an ordered sequence of nodes, presented in document order. Because this view is presented without respect to hierarchy, iterators have methods to move forward and backward, but not to move up and down. Conversely, a TreeWalker maintains the hierarchical relationships of the subtree, allowing navigation of this hierarchy. In general, TreeWalkers are better for tasks in which the structure of the document around selected nodes will be manipulated, while NodeIterators are better for tasks that focus on the content of each selected node.

NodeIterators and TreeWalkers each present a view of a document subtree that may not contain all nodes found in the subtree. In this specification, we refer to this as the logical view to distinguish it from the physical view, which corresponds to the document subtree per se. When an iterator or TreeWalker is created, it may be associated with a NodeFilter, which examines each node and determines whether it should appear in the logical view. In addition, flags may be used to specify which node types should occur in the logical view.

NodeIterators and TreeWalkers are dynamic - the logical view changes to reflect changes made to the underlying document. However, they differ in how they respond to those changes. NodeIterators, which present the nodes sequentially, attempt to maintain their

location relative to a position in that sequence when the sequence's contents change. TreeWalkers, which present the nodes as a filtered tree, maintain their location relative to their current node and remain attached to that node if it is moved to a new context. We will discuss these behaviors in greater detail below.

NodeIterators

A NodeIterator allows the members of a list of nodes to be returned sequentially. In the current DOM interfaces, this list will always consist of the nodes of a subtree, presented in document order. When an iterator is first created, calling its `nextNode()` method returns the first node in the logical view of the subtree; in most cases, this is the root of the subtree. Each successive call advances the NodeIterator through the list, returning the next node available in the logical view. When no more nodes are visible, `nextNode()` returns null.

NodeIterators are created using the `createNodeIterator` method found in the `DocumentTraversal` interface. When a NodeIterator is created, flags can be used to determine which node types will be "visible" and which nodes will be "invisible" while traversing the tree; these flags can be combined using the OR operator. Nodes that are "invisible" are skipped over by the iterator as though they did not exist.

The following code creates an iterator, then calls a function to print the name of each element:

```
NodeIterator iter=
((DocumentTraversal)document).createNodeIterator(
    root, NodeFilter.SHOW_ELEMENT, null);

while (Node n = iter.nextNode())
    printMe(n);
```

Moving Forward and Backward

NodeIterators present nodes as an ordered list, and move forward and backward within

this list. The iterator's position is always either between two nodes, before the first node, or after the last node. When an iterator is first created, the position is set before the first item. The following diagram shows the list view that an iterator might provide for a particular subtree, with the position indicated by an asterisk '*' :

* A B C D E F G H I

Each call to `nextNode()` returns the next node and advances the position. For instance, if we start with the above position, the first call to `nextNode()` returns "A" and advances the iterator:

[A] * B C D E F G H I

The position of a `NodeIterator` can best be described with respect to the last node returned, which we will call the reference node. When an iterator is created, the first node is the reference node, and the iterator is positioned before the reference node. In these diagrams, we use square brackets to indicate the reference node.

A call to `previousNode()` returns the previous node and moves the position backward. For instance, if we start with the `NodeIterator` between "A" and "B", it would return "A" and move to the position shown below:

* [A] B C D E F G H I

If `nextNode()` is called at the end of a list, or `previousNode()` is called at the beginning of a list, it returns null and does not change the position of the iterator. When a `NodeIterator` is first created, the reference node is the first node:

* [A] B C D E F G H I

Robustness

A `NodeIterator` may be active while the data structure it navigates is being edited, so an iterator must behave gracefully in the face of change. Additions and removals in the underlying data structure do not invalidate a `NodeIterator`; in fact, a `NodeIterator` is never invalidated unless its `detach()` method is invoked. To make this possible, the iterator uses

the reference node to maintain its position. The state of an iterator also depends on whether the iterator is positioned before or after the reference node.

If changes to the iterated list do not remove the reference node, they do not affect the state of the NodeIterator. For instance, the iterator's state is not affected by inserting new nodes in the vicinity of the iterator or removing nodes other than the reference node. Suppose we start from the following position:

A B C [D] * E F G H I

Now let's remove "E". The resulting state is:

A B C [D] * F G H I

If a new node is inserted, the NodeIterator stays close to the reference node, so if a node is inserted between "D" and "F", it will occur between the iterator and "F":

A B C [D] * X F G H I

Moving a node is equivalent to a removal followed by an insertion. If we move "I" to the position before "X" the result is:

A B C [D] * I X F G H

If the reference node is removed from the list being iterated over, a different node is selected as the reference node. If the reference node's position is before that of the NodeIterator, which is usually the case after `nextNode()` has been called, the nearest node before the iterator is chosen as the new reference node. Suppose we remove the "D" node, starting from the following state:

A B C [D] * F G H I

The "C" node becomes the new reference node, since it is the nearest node to the NodeIterator that is before the iterator:

A B [C] * F G H I

If the reference node is after the NodeIterator, which is usually the case after previousNode() has been called, the nearest node after the iterator is chosen as the new reference node. Suppose we remove "E", starting from the following state:

A B C D * [E] F G H I

The "F" node becomes the new reference node, since it is the nearest node to the NodeIterator that is after the iterator:

A B C D * [F] G H I

As noted above, moving a node is equivalent to a removal followed by an insertion. Suppose we wish to move the "D" node to the end of the list, starting from the following state:

A B C [D] * F G H I C

The resulting state is as follows:

A B [C] * F G H I D

One special case arises when the reference node is the last node in the list and the reference node is removed. Suppose we remove node "C", starting from the following state:

A B * [C]

According to the rules we have given, the new reference node should be the nearest node after the NodeIterator, but there are no further nodes after "C". The same situation can arise when previousNode() has just returned the first node in the list, which is then removed. Hence: If there is no node in the original direction of the reference node, the nearest node in the opposite direction is selected as the reference node:

A [B] *

If the NodeIterator is positioned within a block of nodes that is removed, the above rules

clearly indicate what is to be done. For instance, suppose "C" is the parent node of "D", "E", and "F", and we remove "C", starting with the following state:

A B C [D] * E F G H I D

The resulting state is as follows:

A [B] * G H I D

Finally, note that removing a NodeIterator's root node from its parent does not alter the list being iterated over, and thus does not change the iterator's state.

Visibility of Nodes

The underlying data structure that is being iterated may contain nodes that are not part of the logical view, and therefore will not be returned by the NodeIterator. If nodes that are to be excluded because of the value of the whatToShow flag, nextNode() returns the next visible node, skipping over the excluded "invisible" nodes. If a NodeFilter is present, it is applied before returning a node; if the filter does not accept the node, the process is repeated until a node is accepted by the filter and is returned. If no visible nodes are encountered, a null is returned and the iterator is positioned at the end of the list. In this case, the reference node is the last node in the list, whether or not it is visible. The same approach is taken, in the opposite direction, for previousNode().

In the following examples, we will use lowercase letters to represent nodes that are in the data structure, but which are not in the logical view. For instance, consider the following list:

A [B] * c d E F G

A call to nextNode() returns E and advances to the following position:

A B c d [E] * F G

Nodes that are not visible may nevertheless be used as reference nodes if a reference node is removed. Suppose node "E" is removed, started from the state given above. The

resulting state is:

A B c [d] * F G

Suppose a new node "X", which is visible, is inserted before "d". The resulting state is:

A B c X [d] * F G

Note that a call to previousNode() now returns node X. It is important not to skip over invisible nodes when the reference node is removed, because there are cases, like the one just given above, where the wrong results will be returned. When "E" was removed, if the new reference node had been "B" rather than "d", calling previousNode() would not return "X".

DYNAMIC DOCUMENTS WITH JAVASCRIPT

4.11 Introduction to dynamic documents

Dynamic HTML is not a new markup language

- A dynamic HTML document is one whose tag attributes, tag contents, or element style properties can be changed after the document has been and is still being displayed by a browser
- We will discuss only W3C standard approaches
- All examples in this chapter, except the last, use the DOM 0 event model and work with both IE6 and NS6
- To make changes in a document, a script must be able to address the elements of the document using the DOM addresses of those elements

4.12 Positioning elements

For DHTML content developers, the most important feature of CSS is the ability to use ordinary CSS style attributes to specify the visibility, size, and position of individual elements of a document. In order to do DHTML programming, it is important to understand how these style attributes work. They are summarized in Table and documented in more detail in the sections that follow.

Table. CSS positioning and visibility attributes

| Attribute(s) | Description |
|------------------|--|
| position | Specifies the type of positioning applied to an element |
| top, left | Specifies the position of the top and left edges of an element |
| bottom,
right | Specifies the position of the bottom and right edges of an element |
| width,
height | Specifies the size of an element |

| | |
|------------|---|
| z-index | Specifies the "stacking order" of an element relative to any overlapping elements; defines a third dimension of element positioning |
| display | Specifies how and whether an element is displayed |
| visibility | Specifies whether an element is visible |
| clip | Defines a "clipping region" for an element; only portions of the element within this region are displayed |
| overflow | Specifies what to do if an element is bigger than the space allotted for it |

The Key to DHTML: The position Attribute

The CSS position attribute specifies the type of positioning applied to an element. The four possible values for this attribute are:

static

This is the default value and specifies that the element is positioned according to the normal flow of document content (for most Western languages, this is left to right and top to bottom.) Statically positioned elements are not DHTML elements and cannot be positioned with the top, left, and other attributes. To use DHTML positioning techniques with a document element, you must first set its position attribute to one of the other three values.

absolute

T
h
i
s

allows you to specify the position of an element relative to its containing element. Absolutely positioned elements are positioned independently of all other elements and are not part of the flow of statically positioned elements.

Fixed An absolutely positioned element is positioned either relative to the `<body>` of the document or, if it is nested within another absolutely positioned element, relative to that element. This is the most commonly used positioning type for DHTML.¹

¹This value allows you to specify an element's position with respect to the browser window. Elements with fixed positioning do not scroll with the rest of the document and thus can be used to achieve frame-like effects. Like absolutely positioned elements, fixed-position elements are independent of all others.

relative

When the position attribute is set to relative, an element is laid out according to the normal flow, and its position is then adjusted relative to its position in the normal flow. The space allocated for the element in the normal document flow remains allocated for it, and the elements on either side of it do not close up to fill in that space, nor are they "pushed away" from the new position of the element. Relative positioning can be useful for some static graphic design purposes, but it is not commonly used for DHTML effects.

Specifying the Position and Size of Elements

Once you have set the position attribute of an element to something other than static, you can specify the position of that element with some combination of the left , top, right, and bottom attributes. The most common positioning technique is to specify the left and top attributes, which specify the distance from the left edge of the containing element (usually the document itself) to the left edge of the element, and the distance from the top edge of the container to the top edge of the element. For example, to place an element 100 pixels from the left and 100 pixels from the top of the document, you can specify CSS styles in a style attribute as follows:

```
<div style="position: absolute; left: 100px; top: 100px;">
```

The containing element relative to which a dynamic element is positioned is not necessarily the same as the containing element within which the element is defined in the document source. Since dynamic elements are not part of normal element flow, their positions are not specified relative to the static container element within which they are defined. Most dynamic elements are positioned relative to the document (the <body> tag) itself. The exception is dynamic elements that are defined within other dynamic elements. In this case, the nested dynamic element is positioned relative to its nearest dynamic ancestor.

Although it is most common to specify the position of the upper-left corner of an element with left and top, you can also use right and bottom to specify the position of the bottom and right edges of an element relative to the bottom and right edges of the containing element. For example, to position an element so that its bottom-right corner is at the

bottom-right of the document (assuming it is not nested within another dynamic element), use the following styles:

```
position: absolute; right: 0px; bottom: 0px;
```

To position an element so that its top edge is 10 pixels from the top of the window and its right edge is 10 pixels from the right of the window, you can use these styles:

```
position: fixed; right: 10px; top: 10px;
```

Note that the right and bottom attributes are newer additions to the CSS standard and are not supported by fourth-generation browsers, as top and left are.

In addition to the position of elements, CSS allows you to specify their size. This is most commonly done by providing values for the width and height style attributes. For example, the following HTML creates an absolutely positioned element with no content. Its width, height, and background-color attributes make it appear as a small blue square:

```
<div style="position: absolute; left: 10px; right: 10px;  
width: 10px; height: 10px; background-color: blue">  
</div>
```

Another way to specify the width of an element is to specify a value for both the left and right attributes. Similarly, you can specify the height of an element by specifying both top and bottom. If you specify a value for left, right, and width, however, the width attribute overrides the right attribute; if the height of an element is over-constrained, height takes priority over bottom.

Bear in mind that it is not necessary to specify the size of every dynamic element. Some elements, such as images, have an intrinsic size. Furthermore, for dynamic elements that contain text or other flowed content, it is often sufficient to specify the desired width of the element and allow the height to be determined automatically by the layout of the element's content.

In the previous positioning examples, values for the position and size attributes were specified with the suffix "px". This stands for pixels. The CSS standard allows measurements to be done in a number of other units, including inches ("in"), centimeters

("cm"), points ("pt"), and ems ("em" -- a measure of the line height for the current font). Pixel units are most commonly used with DHTML programming. Note that the CSS standard requires a unit to be specified. Some browsers may assume pixels if you omit the unit specification, but you should not rely on this behavior.

Instead of specifying absolute positions and sizes using the units shown above, CSS also allows you to specify the position and size of an element as a percentage of the size of the containing element. For example, the following HTML creates an empty element with a black border that is half as wide and half as high as the containing element (or the browser window) and centered within that element:

```
<div style="position: absolute; left: 25%; top: 25%; width: 50%; height: 50%;  
border: 2px solid black">  
</div>
```

Element size and position details

It is important to understand some details about how the left , right, width, top, bottom, and height attributes work. First, width and height specify the size of an element's content area only; they do not include any additional space required for the element's padding, border, or margins. To determine the full onscreen size of an element with a border, you must add the left and right padding and left and right border widths to the element width, and you must add the top and bottom padding and top and bottom border widths to the element's height.

Since width and height specify the element content area only, you might think that left and top (and right and bottom) would be measured relative to the content area of the containing element. In fact, the CSS standard specifies that these values are measured relative to the outside edge of the containing element's padding (which is the same as the inside edge of the element's border).

Let's consider an example to make this clearer. Suppose you've created a dynamically positioned container element that has 10 pixels of padding all the way around its content area and a 5 pixel border all the way around the padding. Now suppose you dynamically position a child element inside this container. If you set the left attribute of the child to "0

px", you'll discover that the child is positioned with its left edge right up against the inner edge of the container's border. With this setting, the child overlaps the container's padding, which presumably was supposed to remain empty (since that is the purpose of padding). If you want to position the child element in the upper left corner of the container's content area, you should set both the left and top attributes to "10px". Figure helps to clarify this.

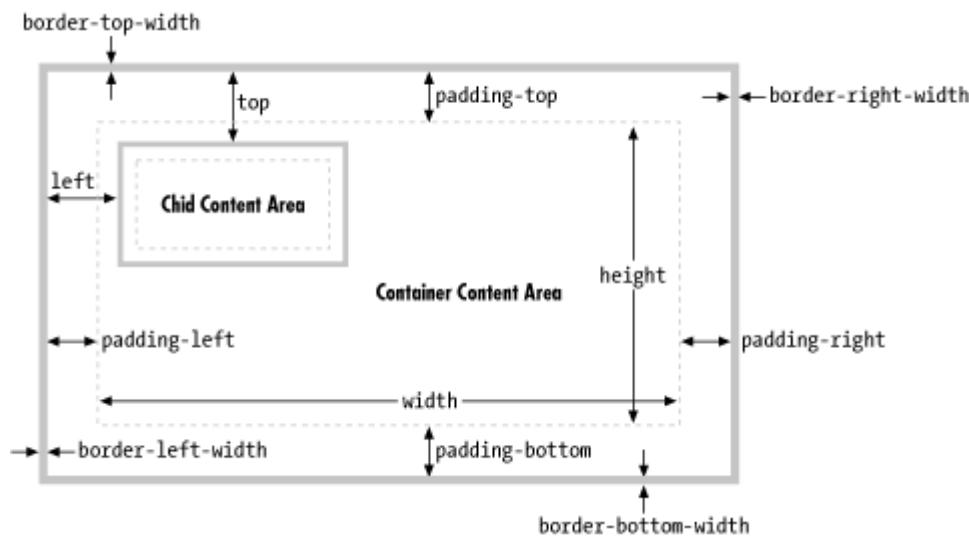


Figure Dynamically positioned container and child elements with some CSS attributes

Now that you understand that width and height specify the size of an element's content area only and that the left, top, right, and bottom attributes are measured relative to the containing element's padding, there is one more detail you must be aware of: Internet Explorer Versions 4 through 5.5 for Windows (but not IE 5 for the Mac) implement the width and height attributes incorrectly and include an element's border and padding (but not its margins). For example, if you set the width of an element to 100 pixels and place a 10-pixel margin and a 5-pixel border on the left and right, the content area of the element ends up being only 70 pixels wide in these buggy versions of Internet Explorer.

In IE 6, the CSS position and size attributes work correctly when the browser is in standards mode and incorrectly (but compatibly with earlier versions) when the browser is in compatibility mode. Standards mode, and hence correct implementation of the CSS

"box model," is triggered by the presence of a <!DOCTYPE> tag at the start of the document, declaring that the document adheres to the HTML 4.0 (or later) standard or some version of the XHTML standards. For example, any of the following three HTML document type declarations cause IE 6 to display documents in standards mode:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN">
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Strict//EN">
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN"
  "http://www.w3.org/TR/html4/loose.dtd">
```

Netscape 6 and the Mozilla browser handle the width and height attributes correctly. But these browsers also have standards and compatibility modes, just as IE does. The absence of a <!DOCTYPE> declaration puts the Netscape browser in quirks mode, in which it mimics certain (relatively minor) nonstandard layout behaviors of Netscape 4. The presence of <!DOCTYPE> causes the browser to break compatibility with Netscape 4 and correctly implement the standards.

4.13 Moving elements

```
id = window.setInterval("animate()", 100);
```

A rather rare, but still used, DHTML scenario is not only positioning an element, but also moving and therefore animating an element. To do so, window.setTimeout() and window.setInterval() come in handy. The following code animates an ad banner diagonally over the page. The only potential issue is how to animate the position. For the Internet Explorer properties (posLeft, posTop), just adding a value suffices. For left and top, you first have to determine the old position and then add a value to it. The JavaScript function parseInt() extracts the numeric content from a string like "123px". However, parseInt() returns NaN if no value is found in left or top. Therefore, the following helper function takes care of this situation; in this case, 0 is returned instead:

```
function myParseInt(s) {
  var ret = parseInt(s);
  return (isNaN(ret) ? 0 : ret);
```

```
}
```

Then, the following code animates the banner and stops after 50 iterations:

Animating an Element (animate.html; excerpt)

```
<script language="JavaScript"
  type="text/javascript">
var nr = 0; var id =
null; function
animate() { nr++;
if (nr > 50) { window.clearInterval(id);
  document.getElementById("element").style.visibility = "hidden";
} else {
  var el = document.getElementById("element");
  el.style.left =
    (myParseInt(el.style.left) + 5) + "px";
  el.style.posLeft += 5;
  el.style.top =
    (myParseInt(el.style.top) + 5) + "px";
  el.style.posTop += 5;
}
}
window.onload = function() {
  id = window.setInterval("animate()", 100);
};
</script>
```

```
<h1>My Portal</h1>
<div id="element" style="position: absolute;
background-color: #eee; border: 1px solid">
  JavaScript Phrasebook
</div>
```

4.14 Element visibility

The visibility property of an element controls whether it is displayed

- The values are visible and hidden
-

- Suppose we want to toggle between hidden and visible, and the element's DOM address is dom

```

if (dom.visibility == "visible"
    dom.visibility = "hidden";
else
    dom.visibility = "visible";
--> SHOW showHide.html

```

Changing colors and fonts

In order to change text colors, you will need two things:

1. A command to change the text.

2. A color (hex) code.

Section 1: Changing Full-Page Text Colors

You have the ability to change full-page text colors over four levels:

<TEXT="#" -- This denotes the full-page text color.

<LINK="#" -- This denotes the color of the links on your page.

<ALINK="#" -- This denotes the color the link will flash when clicked upon.

<VLINK="#" -- This denotes the colors of the links after they have been visited.

These commands come right after the <TITLE> commands. Again, in that position they affect everything on the page. **Also...** place them all together inside the same command along with any background commands. Something like this:

```
< BODY BGCOLOR="#" TEXT="#" LINK="#" VLINK="#">
<VLINK="#FFFFFF">
```

Section 2: Changing Specific Word Color

But I only want to change one word's color!

You'll use a color (hex) code to do the trick. Follow this formula:

```
<FONT COLOR="######">text text text text text</FONT>
```

It's a pain in the you-know-where, but it gets the job done. It works with all H commands and text size commands. Basically, if it's text, it will work.

Using Cascading **Style Sheets** (CSS) to Change Text Colors

There isn't enough space to fully describe what CSS is capable of in this article, but we have several articles here that can get you up to speed in no time! For a great tutorial on using CSS to change color properties, check out this article by Vincent Wright.

A quick intro to CSS is in order, so let's describe it a bit. CSS is used to define different elements on your web page. These elements include text colors, link colors, page background, tables, forms--just about every aspect of the style of the web page. You can use CSS inline, much like the HTML above, or you can, more preferably, include the style sheet within the HEAD tags on your page, as in this example:

```
<STYLE type=text/css>
A:link {
    COLOR: red /*The color of the link*/
}

A:visited {
    COLOR: #800080 /*The color of the visited link*/
}

A:hover {
    COLOR: green /*The color of the mouseover or 'hover' link*/
}

BODY { COLOR: #800080 /*The color of all the other text within the body of the page*/
}

</STYLE>
```

Alternately, you can include the CSS that is between the STYLE tags above, and save it in a file that you could call "basic.css" which would be placed in the root directory of your website. You would then refer to that style sheet by using a link that goes between

the HEAD tags in your web page, like this:

```
<link type="text/css" rel="stylesheet" href="basic.css">
```

As you can see in the example above, you can refer to the colors using traditional color names, or hex codes as described above.

The use of CSS is vastly superior to using inline FONT tags and such, as it separates the content of your site from the style of your site, simplifying the process as you create more pages or change the style of elements. If you are using an external style sheet, you can make the change once in the style sheet, and it will be applied to your entire site. If you choose to include the style sheet itself within the HEAD tags as shown above, then you will have to make those changes on every page on your site.

CSS is such a useful tool in your [web developer](#) arsenal, you should definitely take the time to read more about it in our CSS Tutorials section.

4.15 Dynamic content

In the early days of the Web, once a Web page was loaded into a client's browser, changing the information displayed on the page required another call to the Web server.

The user interacted with the Web page by submitting forms or clicking on links to other Web pages, and the Web server delivered a new page to the client.

Any customization of information - such as building a Web page on the fly from a template - required that the Web server spent additional time processing the page request.

In short dynamic content is about changing the content of a Web page by inserting and deleting elements or the content inside elements before, or after, a Web page has been loaded into a client's browser.

UNIT - 5

XML

5.1 Introduction

SGML is a meta-markup language is a language for defining markup language it can describe a wide variety of document types.

- _ Developed in the early 1980s; In 1986 SGML was approved by ISO std.
- _ HTML was developed using SGML in the early 1990s - specifically for Web documents.
- _ Two problems with HTML:
 1. HTML is defined to describe the general form and layout of information without considering its meaning.
 2. Fixed set of tags and attributes. Given tags must fit every kind of document. No way to find particular information
 3. There are no restrictions on arrangement or order of tag appearance in document. For example, an opening tag can appear in the content of an element, but its corresponding closing tag can appear after the end of the element in which it is nested.

Eg : Now is the time

- _ One solution to the first problems is to allow for group of users with common needs to define their own tags and attributes and then use the SGML standard to define a new markup language to meet those needs. Each application area would have its own markup language.

- _ Use SGML to define a new markup language to meet those needs

- _ Problem with using SGML:

1. It's too large and complex to use and it is very difficult to build a parser for it. SGML includes a large number of capabilities that are only rarely used.
2. A program capable of SGML documents would be very large and costly to develop.
3. SGML requires that a formal definition be provided with each new markup language.

So having area-specific markup language is a good idea, basing them on SGML is not.

- _ A better solution: Define a simplified version of SGML and allow users to define their own markup languages based on it. XML was designed to be that simplified version of

SGML.

- _ XML is not a replacement for HTML . Infact two have different goals
- _ HTML is a markup language used to describe the layout of any kind of information
- _ XML is a meta-markup language that provides framework for defining specialized markup languages

- _ Instead of creating text file like

```
<html>  
<head><title>name</title></head>.....
```

Syntax

```
<name>  
<first> nandini </first>  
<last> sidnal </last>  
</name>
```

XML is much larger then text file but makes easier to write software that accesses the information by giving structure to data.

- _ XML is a very simple and universal way of storing and transferring any textual kind
- _ XML does not predefine any tags
- XML tag and its content, together with closing tag _ element
- _ XML has no hidden specifications
- _ XML based markup language as _ tag set
- _ Document that uses XML based markup language _ XML document
- _ An XML processor is a program that parses XML documents and provides the parts to an application
- _ Both IE7 and FX2 support basic XML XML is a meta language for describing mark-up languages. It provides a facility to define tags and the structural relationship between them

What is XML?

- XML stands for EXtensible Markup Language
- XML is a markup language much like HTML
- XML was designed to carry data, not to display data
- XML tags are not predefined. You must define your own tags
- XML is designed to be self-descriptive

- XML is a W3C Recommendation

XML is not a replacement for HTML.

XML syntax rules:

The syntax of XML can be thought of at two distinct levels.

1. There is the general low-level syntax of XML that imposes its rules on all XML documents.

2. The other syntactic level is specified by either document type definitions or XML schemas. These two kinds of specifications impose structural syntactic rules on documents written with specific XML tag sets.

_ DTDs and XML schema specify the set of tags and attributes that can appear in particular document or collection of documents, and also the orders and various arrangements in which they can appear.

DTD's and XML schema can be used to define a XML markup language.

XML document can include several different kinds of statements.

- Data elements
- Markup declarations - instructions to XML parser
- Processing instructions – instructions for an applications program that will process the data described in the document. The most common of these are the data elements of the document. XML document may also include markup declarations, which are instructions to the XML parser, and processing instructions, which are instructions for an application program that will process the data described in the document.

All XML document must begin with XML declaration. It identifies the document as being XML and provides the version no. of the XML standard being used. It will also include encoding standard. It is a first line of the XHTML document.

-- This is a comment -->

XML names must begin with a letter or underscore and can include digits, hyphens, and periods.

XML names are case sensitive. , the tag <Letter> is different from the tag <letter>.

There is no length limitation for names.

[redacted] pace is Preserved in XML

HTML truncates multiple white-space characters to one single white-space:

HTML: Hello my name is Tove

Output: Hello my name is Tove.

With XML, the white-space in a document is not truncated.

Every XML document defines a single root element, whose opening tag must appear on the first line of XML code.

ments must be nested inside the root element. The root element of every XHTML document is html.

<element_name/> --- no content

<p>This is a paragraph

<p>This is another paragraph.

must have a closing tag:

<p>This is a paragraph</p>

<p>This is another paragraph</p>



<Message>This is incorrect</message>

<message>This is correct</message>

In HTML, you might see improperly nested elements:

<i>This text is bold and italic</i>

In XML, all elements **must** be properly nested within each other:

<i>This text is bold and italic</i>

In the example above, "Properly nested" simply means that since the *<i>* element is opened inside the ** element, it must be closed inside the ** element.

XML Documents Must Have a Root Element

It contain one element that is the **parent** of all other elements.

This element is called the **root** element.

```
<root>
<child>
<subchild>.....</subchild>
</child>
</root>
```

XML tags can have attributes, which are specified with name/value assignments. XML Attribute Values must be enclosed with single or double quotation marks. XML document that strictly adheres to these syntax rule is considered as well formed. An XML document that follows all of these rules is well formed

Example :

```
<?xml version = "1.0" encoding = "utf-8"?>
<ad>
<year>1960</year>
<make>Cessna</make>
<model>Centurian</model>
<color>Yellow with white trim</color>
<location>
<city>Gulfport</city>
<state>Mississippi</state>
</location>
</ad>
```

None of this tag in the document is defined in XHTML-all are designed for the specific content of the document. When designing an XML document, the designer is often faced with the choice between adding a new attribute to an element or defining a nested element.

- o In some cases there is no choices.
- o In other cases, it may not matter whether an attribute or a nested element is used.
- o Nested tags are used,
 - when tags might need to grow in structural complexity in the future
 - if the data is subdata of the parent element's content
 - if the data has substructure of its own
- o Attribute is used ,
 - For identifying numbers/names of element
 - If data in question is one value from a given possibilities
 - The attribute is used if there is no substructure

```
<!-- A tag with one attribute -->
<patient name = "Maggie Dee Magpie">
...
</patient>

<!-- A tag with one nested tag -->
<patient>
<name> Maggie Dee Magpie </name>
...
</patient>

<!-- A tag with one nested tag, which contains
three nested tags -->
<patient>
<name>
<first> Maggie </first>
<middle> Dee </middle>
<last> Magpie </last>
</name>
...
</patient>
```

Here third one is a better choice because it provides easy access to all of the parts of data.

5.3 Document structure

XML document often uses two auxiliary files:

1. It specifies tag set and syntactic structural rules.
2. It contain the style sheet to describe how the content of the document to be printed.

XML documents (and HTML documents) are made up by the following building blocks:

Elements, Tags, Attributes, Entities, PCDATA, and CDATA

Elements

Elements are the main building blocks of both XML and HTML documents.

Elements can contain text, other elements, or be empty.

Tags

Tags are used to markup elements.

A starting tag like <element_name> mark up the beginning of an element, and an ending tag like

</element_name> mark up the end of an element.

Examples:

A body element: <body>body text in between</body>.

A message element: <message>some message in between</message>

Attributes

Attributes provide extra information about elements.

Attributes are placed inside the start tag of an element. Attributes come in name/value pairs. The following "img" element has an additional information about a source file:

 The name of the element is "img". The name of the attribute is "src".

The value of the attribute is "computer.gif". Since the element itself is empty it is closed by a "
/".

PCDATA

PCDATA means parsed character data.

Think of character data as the text found between the start tag and the end tag of an XML element.

PCDATA is text that will be parsed by a parser. Tags inside the text will be treated as markup and entities will be expanded.

CData

CData also means character data. CData is text that will NOT be parsed by a parser. Tags inside the text will NOT be treated as markup and entities will not be expanded.

Most of you know the HTML entity reference: " " that is used to insert an extra space in an HTML document. Entities are expanded when a document is parsed by an XML parser.

_ An XML document often uses two auxiliary files:

- One to specify the structural syntactic rules (DTD / XML schema)
- One to provide a style specification (CSS /XSLT Style Sheets)

Entities

An XML document has a single root element, but often consists of one or more entities

An XML document consist of one or more entities that are logically related collection of information,

Entities range from a single special character to a book chapter

- An XML document has one document entity
- * All other entities are referenced in the document entity

Reasons to break a document into multiple entities.

1. Good to define a Large documents as a smaller no. of parts easier to manage .
2. If the same data appears in more than one place in the document, defining it as a entity allows any no. of references to a single copy of data.
3. Many documents include information that cannot be represented as text, such as images. Such information units are usually stored as binary data. Binary entities can only be referenced in the document entities

Entity names:

- No length limitation
- Must begin with a letter, a dash, or a colon
- Can include letters, digits, periods, dashes, underscores, or colons

_ A reference to an entity has the form name with prepended ampersand and appended semicolon: &entity_name; Eg. &apple_image;

- _ One common use of entities is for special characters that may be used for markup delimiters to appear themselves in the document.

These are predefined (as in XHTML):

Character	Entity	Meaning
<	<	Less than
>	>	Greater than
&	&	Ampersand
“	"	Double quote
‘	&apos	Single quote

- _ If several predefined entities must appear near each other in a document, it is better to avoid using entity references. Character data section can be used. The content of a character data section is not parsed by the XML parser, so it can include any tags.

- _ The form of a character data section is as follows: <![CDATA[content]]> // no tags can be used since it is not parsed. For example, instead of Start >>>> HERE <<<<

use

```
<![CDATA[Start >>> HERE <<<]]>
```

The opening keyword of a character data section is not just CDATA, it is in effect [CDATA[. There can be any spaces between [and C or between A and].

- _ Content of Character data section is not parsed by parser. For example the content of the line <![CDATA[The form of a tag is <tag name>]]> is as follows
The form of a tag is <tag name>

5.4 Document Type definitions

A DTD is a set of structural rules called declarations. These rules specify a set of elements, along with how and where they can appear in a document

- Purpose: provide a standard form for a collection of XML documents and define a markup language for them.
- DTD provide entity definition.

- With DTD, application development would be simpler.
- Not all XML documents have or need a DTD

1 External style sheets are used to impose a uniform style over a collection of documents.

1 When are DTDs used?

When same tag set definition are used by collection of documents , collection of users and documents must have a consistent and uniform structure.

1 A document can be tested against a DTD to determine whether it conforms to the rules the DTD describes.

1 Application programs that processes the data in the collection of XML documents can be written to assume the particular document form.

1 Without such structural restrictions, developing such applications would be difficult. If not impossible.

1 The DTD for a document can be internal (embedded in XML document) or external(separate file)- can be used with more than one document.

1 DTD with incorrect/inappropriate declaration can have wide-spread consequences.

1 DTD declarations have the form: <!keyword ... >

There are four possible declaration keywords:

ELEMENT, ATTLIST, ENTITY, and NOTATION

1. Declaring Elements:

- Element declarations are similar to BNF(CFG)(used to define syntactic structure of Programming language) here DTD describes syntactic structure of particular set of doc so its rules are similar to BNF.
- An element declaration specifies the name of an element and its structure
- If the element is a leaf node of the document tree, its structure is in terms of characters
- If it is an internal node, its structure is a list of children elements (either leaf or internal nodes)

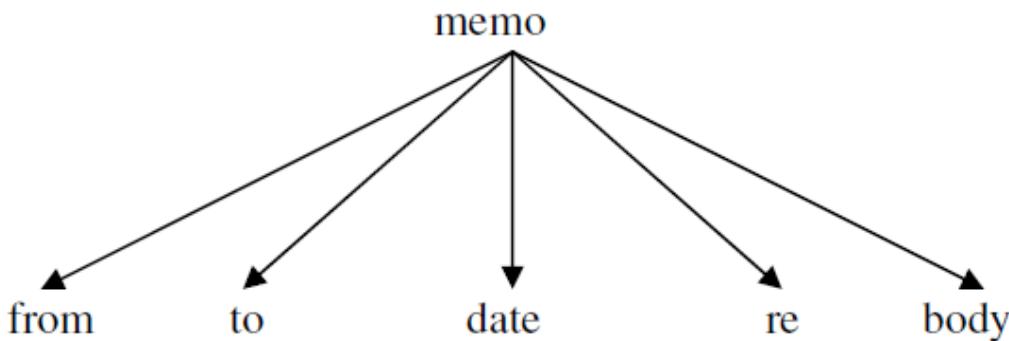
- General form:

```
<!ELEMENT element_name(list of child names)>
```

e.g.,

```
<!ELEMENT memo (from, to, date, re, body)>
```

This element structure can describe the document tree structure shown below.



- Child elements can have modifiers,

– + -> One or more occurrences

– * -> Zero or more occurrences

– ? ->Zero or one occurrences

Ex: consider below DTD declaration

```
<!ELEMENT person (parent+, age, spouse?, sibling*)>
```

– One or more parent elements

– One age element

– Possible a spouse element.

– Zero or more sibling element.

- Leaf nodes specify data types of content of their parent nodes which are elements

1. PCDATA (parsable character data)

2. EMPTY (no content)

3. ANY (can have any content)

Example of a leaf declaration:

```
<!ELEMENT name (#PCDATA)>
```

2. Declaring Attributes:

- Attributes are declared separately from the element declarations
- General form:

```
<!ATTLIST element_name attribute_name attribute_type [default _value]>
```

More than one attribute

```
<!ATTLIST element_name attribute_name1 attribute_type default_value_1  
attribute_name2 attribute_type default_value_2
```

```
...>
```

_ Attribute type :There are ten different types, but we will consider only CDATA

_ Possible Default value for attributes:

Value - value ,which is used if none is specified

#Fixed value - value ,which every element have and can't be changed

Required - no default value is given ,every instance must specify a value

#Implied - no default value is given ,the value may or may not be specified

Example :

```
<!ATTLIST car doors CDATA "4">  
<!ATTLIST car engine_type CDATA #REQUIRED>  
<!ATTLIST car make CDATA #FIXED "Ford">  
<!ATTLIST car price CDATA #IMPLIED>  
<car doors = "2" engine_type = "V8">  
...  
</car>
```

Declaring Entities :

Two kinds:

- A general entity can be referenced anywhere in the content of an XML document

Ex: Predefined entities are all general entities.

- A parameter entity can be referenced only in DTD.
- General Form of entity declaration.

```
<!ENTITY [%] entity_name "entity_value">
```

% when present it specifies declaration parameter entity

Example :

```
<!ENTITY jfk "John Fitzgerald Kennedy">
```

_ A reference above declared entity: &jfk;

- If the entity value is longer than a line, define it in a separate file (an external text entity)

_ General Form of external entity declaration

```
<!ENTITY entity_name SYSTEM -file_location">
```

SYSTEM specifies that the definition of the entity is in a different file.

- Example for parameter entity

```
<!ENTITY %pat -(USN, Name)||>
```

```
<!ELEMENT student %pat; >
```

4. Sample DTD:

```
<?xml version = "1.0" encoding = "utf-8"?>
```

```
<!-- planes.dtd - a document type definition for the planes.xml document, which specifies  
a list of used airplanes for sale -->
```

```
<!ELEMENT planes_for_sale (ad+)>
```

```
<!ELEMENT ad (year, make, model, color, description, price?, seller, location)>
```

```
<!ELEMENT year (#PCDATA)>
```

```
<!ELEMENT make (#PCDATA)>
```

```
<!ELEMENT model (#PCDATA)>
```

```
<!ELEMENT color (#PCDATA)>
```

```
<!ELEMENT description (#PCDATA)>
```

```
<!ELEMENT price (#PCDATA)>
```

```
<!ELEMENT seller (#PCDATA)>
```

```
<!ELEMENT location (city, state)>
```

```
<!ELEMENT city (#PCDATA)>
```

```
<!ELEMENT state (#PCDATA)>
```

```
<!ATTLIST seller phone CDATA #REQUIRED>
```

```
<!ATTLIST seller email CDATA #IMPLIED>
```

```
<!ENTITY c "Cessna">
```

```
<!ENTITY p "Piper">
```

```
<!ENTITY b "Beechcraft">
```

5. Internal and External DTD's:

- Internal DTDs

```
<!DOCTYPE planes [  
    <!-- The DTD for planes -->  
]>
```

- External DTDs

```
<!DOCTYPE XML_doc_root_name SYSTEM  
    -DTD_file_name|>
```

For examples,

```
<!DOCTYPE planes_for_sale SYSTEM  
    -planes.dtd|>  
<?xml version = "1.0" encoding = "utf-8"?>  
<!-- planes.xml - A document that lists ads for used airplanes -->  
<!DOCTYPE planes_for_sale SYSTEM "planes.dtd">  
<planes_for_sale>  
<ad>  
<year> 1977 </year>  
<make> &c; </make>  
<model> Skyhawk </model>  
<color> Light blue and white </color>  
<description> New paint, nearly new interior,  
685 hours SMOH, full IFR King avionics </description>  
<price> 23,495 </price>  
<seller phone = "555-222-3333"> Skyway Aircraft </seller>  
<location>  
<city> Rapid City, </city>  
<state> South Dakota </state>  
</location>  
</ad>  
</planes_for_sale>
```

5.5 Namespaces

1 XML provides benefits over bit formats and can now create well formed XML doc.

But when applications become complex we need to combine elements from various doc types into one XML doc. This causes Pb?

1 Two documents will have elements with same name but with different meanings and semantics.

1 Namespaces – a means by which you can differentiate elements and attributes of different XML document types from each other when combining them together into other documents , or even when processing multiple documents simultaneously.

1 Why do we need Namespaces?

1 XML allows users to create their tags, naming collisions can occur For ex: element title

– <title>Resume of John</title>

– <title>Sir</title><Fname>John</Fname>

1 Namespaces provide a means for user to prevent naming collisions Element title

– <xhtml:title>Resume of John</xhtml:title>

– <person:title>Sir</person:title><Fname>John</Fname> xhtml and person --- two namespaces

1 Namespace is collection of elements and attribute names used in XML documents. It has a form of a URI

1 A Namespace for elements and attributes of the hierarchy rooted at particular element is declared as the value of the attribute xmlns. Namespace declaration for an element is

<element_name xmlns[:prefix] = URI>

- The square bracket indicates that what is within them is optional.

- prefix[optional] specify name to be attached to names in the declared namespace

1 Two reasons for prefix :

1. Shorthand for URI // URI is too long to be typed on every occurrence of every name from the namespace.

2. URI may includes characters that are illegal in XML

- 1 Usually the element for which a namespace is declared is usually the root of a document.

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

This declares the default namespace, from which names appear without prefixes.

As an example of a prefixed namespace declaration, consider

```
<birds xmlns:bd = http://www.audubon.org/names/species>
```

Within the birds element, including all of its children elements, the names from the namespace must be prefixed with bd, as in the following.

```
<bd:lark>
```

If an element has more than one namespace declaration, then

```
<birds xmlns:bd = -http://www.audubon.org/names/species||>
```

```
xmlns : html = -http://www.w3.org/1999/xhtml||>
```

Here we have added the standard XHTML namespace to the birds element. One namespace declaration in an element can be used to declare a default namespace. This can be done by not including the prefix in the declaration. The names from the default by omitting the prefix.

Consider the example in which two namespaces are declared.

The first is declared to be the default namespaces.

The second defines the prefix, cap.

```
<states>
```

```
xmlns = http://www.states-info.org/states
```

```
xmlns:cap = http://www.states-info.org/state-capitals
```

```
<state>
```

```
<name> South Dakota </name>
```

```
<population> 75689 </population>
```

```
<capital>
```

```

<cap:name> Pierre </cap:name>
<cap:population>12429</cap:population>
</capital>
</state>
</states>

```

Each state element has name and population elements for both namespaces.

5.6 XML schemas

1 A schema is any type of model document that defines the structure of something, such as databases structure or documents. Here something is XML doc. Actually DTDs are a type of schema.

1 An XML schema is an XML document so it can be parsed with an XML parser.

1 The term XML schema is used to refer to specify W3C XML schema technology.

1 W3C XML Schemas like DTD allow you to describe the structure for an XML doc.

1 DTDs have several disadvantages

- Syntax is different from XML - cannot be parsed with an XML parser
- It is confusing to deal with two different syntactic forms
- DTDs do not allow restriction on the form of data that can be content of element ex: `<quantity>5</quantity>` and `<quantity>5</quantity>` are valid DTD can only specifies that could be anything. Eg time No datatype for integers all are treated as texts.

1 XML Schemas is one of the alternatives to DTD

- It is XML document, so it can be parsed with XML parser
- It also provides far more control over data types than do DTDs
- User can define new types with constraints on existing data types

1. Schema Fundamentals:

- Schema are related idea of class and an object in an OOP language

D Schema class definition

D XML document confirming to schema structure 1 Object

- Schemas have two primary purposes

D Specify the structure of its instance XML documents

D Specify the data type of every element & attribute of its instance XML documents

2. Defining a schema:

Schemas are written from a namespace(schema of schemas):

<http://www.w3.org/2001/XMLSchema> element, schema, sequence and string are some names from this namespace

1 Every XML schema has a single root, schema.

- The schema element must specify the namespace for the schema of schemas from which the schema's elements and its attributes will be drawn.
- It often specifies a prefix that will be used for the names in the schema. This namespace spec appears as

`xmlns:xsd = http://www.w3.org/2001/XMLSchema`

1 Every XML schema itself defines a tag set like DTD, which must be named with the targetNamespace attribute of schema element. The target namespace is specified by assigning a name space to the target namespace attribute as the following:
`targetNamespace = http://cs.uccs.edu/planeSchema`

Every top-level element places its name in the target namespace. If we want to include nested elements, we must set the elementFormDefault attribute to qualified.
`elementFormDefault = qualified`.

1 The default namespace which is source of the unprefixed names in the schema is given with another xmlns specification `xmlns = "http://cs.uccs.edu/planeSchema"`

1 A complete example of a schema element:

```
<xsd:schema
```

```
<!-- Namespace for the schema itself -->
```

```
xmlns:xsd = http://www.w3.org/2001/XMLSchema
```

```

<!-- Namespace where elements defined here will be placed -->
targetNamespace = -http://cs.uccs.edu/planeSchema||

<!-- Default namespace for this document -->
xmlns = -http://cs.uccs.edu/planeSchema||

<!-- Specify non-top-level elements to be in the target namespace-->
elementFormDefault = "qualified" >

```

Defining a schema instance:

- An instance of schema must specify the namespaces it uses
 - These are given as attribute assignments in the tag for its root element
1. Define the default namespace

<planes

xmlns = http://cs.uccs.edu/planesScema

...>

2. It is root element of an instance document is for the schemaLocation attribute.

Specify the standard namespace for instances (XMLSchema-instance) xmlns:xsi
=-http://www.w3.org/2001/XMLSchema-instance"

3. Specify location where the default namespace is defined, using the schemaLocation attribute, which is assigned two values namespace and filename.

xsi:schemaLocation ="http://cs.uccs.edu/planeSchema planes.xsd" >

4. Schema Data types: Two categories of data types

1.Simple (strings only, no attributes and no nested elements)

2. Complex (can have attributes and nested elements)

- XML Schema defines 44 data types
- Primitive: string, Boolean, float, ...
- Derived: byte, decimal, positiveInteger, ...

- User-defined (derived) data types – specify constraints on an existing type (then called as base type)
 - Constraints are given in terms of facets of the base type

Ex: interget data type has *8 facets :totalDigits, maxInclusive....

- 1 Both simple and complex types can be either named or anonymous
- 1 DTDs define global elements (context of reference is irrelevant). But context of reference is essential in XML schema
- 1 Data declarations in an XML schema can be
 1. Local ,which appears inside an element that is a child of schema
 2. Global, which appears as a child of schema

5. Defining a simple type:

- Use the element tag and set the name and type attributes

```
<xsd:element name = "bird" type = "xsd:string"/>
```

D The instance could be :

```
<bird> Yellow-bellied sap sucker </bird>
```

- An element can be given default value using default attribute

```
<xsd:element name = "bird" type = "xsd:string" default="Eagle" />
```

- An element can have constant value, using fixed attribute

```
<xsd:element name = "bird" type = "xsd:string" fixed="Eagle" />
```

Declaring User-Defined Types:

- User-Define type is described in a simpleType element, using facets
- facets must be specified in the content of restriction element
- facets values are specified with the value attribute

For example, the following declares a user-defined type , firstName

```
<xsd:simpleType name = "firstName" >
<xsd:restriction base = "xsd:string" >
<xsd:maxLength value = "20" />
```

```

</xsd:restriction>
</xsd:simpleType>
<xsd:simpleType name = -phoneNumber" >
<xsd:restriction base = "xsd:decimal" >
<xsd:precision value = -40" />
</xsd:restriction>
</xsd:simpleType>

```

6. Declaring Complex Types:

- There are several categories of complex types, but we discuss just one, element-only elements
- Element-only elements are defined with the complex Type element
- Use the sequence tag for nested elements that must be in a particular order
- Use the all tag if the order is not important
- Nested elements can include attributes that give the allowed number of occurrences (minOccurs, maxOccurs, unbounded)
- For ex:

```

<xsd:complexType name = "sports_car" >
<xsd:sequence>
<xsd:element name = "make-type = "xsd:string" />
<xsd:element name = "model-type = "xsd:string" />
<xsd:element name = "engine-type = "xsd:string" />
<xsd:element name = "year-type = "xsd:string" />
</xsd:sequence>
</xsd:complexType>

```

7. Validating Instances of Schemas:

- An XML schema provides a definition of a category of XML documents.

- However, developing a schema is of limited value unless there is some mechanical way to determine whether a given XML instance document confirms to the schema.
- Several XML schema validation tools are available eg. xsv(XML schema validator) This can be used to validate online.
- Output of xsv is an XML document. When run from command line output appears without being formated.

Output of xsv when run on planes.xml

```
<?XML version = _1.0‘ encoding = _utf-8?>
<xsv docElt = {_ http://cs.uccs.edu/planesSchema} planes‘
instanceAccessed = _true‘
instanceErrors = _0‘
schemaErrors = _0‘
schemaLocs = _http://cs.uccs.edu/planesSchema->planes.xsd‘
Target = _file: /c:/wbook2/xml/planes.xml‘
Validation = _strict‘
Version = _XSV 1.197/1.101 of 2001/07/07 12:01:19‘
xmlns=_http://www.w3.org/2000/05.xsv‘>
<importAttempt URI=_file:/c:wbook2/xml/planes.xsd‘
namespace = _http://cs.uccs.edu/planesSchema‘
outcome = _success‘ />
</xsv>
```

If schema is not in the correct format, the validator will report that it could not find the specified schema.

5.7 Displaying RAW XML Documents

- 1 An XML enabled browser or any other system that can deal with XML documents cannot possibly know how to format the tags defined in the doc.
- 1 Without a style sheet that defines presentation styles for the doc tags the XML doc can not be displayed in a formatted manner.
- 1 Some browsers like FX2 have default style sheets that are used when style sheets are not defined.
- 1 Eg of planes.xml document.

5.8 Displaying XML Documents with CSS

- 1 Style sheet information can be provided to the browser for an xml document in two ways.
 - First, a CSS file that has style information for the elements in the XML document can be developed.
 - Second the XSLT style sheet technology can be used..
- 1 Using CSS is effective, XSLT provides far more power over the appearance of the documents display.
- 1 A C SS style sheet for an XML document is just a list of its tags and associated styles
 - 1 The connection of an XML document and its style sheet is made through an xmlstylesheet processing instruction
 - 1 Display— used to specify whether an element is to be displayed inline or in a separate block.

```
<?xml-stylesheet type = "text/css— href = -planes.css"?>
```

For example: planes.css

```
<!-- planes.css - a style sheet for the planes.xml document -->  
ad { display: block; margin-top: 15px; color: blue; }  
year, make, model { color: red; font-size: 16pt; }
```

```
color {display: block; margin-left: 20px; font-size: 12pt;}  
description {display: block; margin-left: 20px; font-size: 12pt;}  
seller { display: block; margin-left: 15px; font-size: 14pt; }  
location {display: block; margin-left: 40px; }  
city {font-size: 12pt; }  
state {font-size: 12pt; }  
<?xml version = "1.0" encoding = "utf-8"?>  
<!-- planes.xml - A document that lists ads for used airplanes -->  
<planes_for_sale>  
<ad>  
<year> 1977 </year>  
<make> Cessana </make>  
<model> Skyhawk </model>  
<color> Light blue and white </color>  
<description> New interior  
</description>  
<seller phone = "555-222-3333">  
Skyway Aircraft </seller>  
<location>  
<city> Rapid City, </city>  
<state> South Dakota </state>  
</location>  
</ad>  
</planes_for_sale>
```

With planes.css the display of planes.xml as following: 1977 Cessana Skyhawk Light blue and white New interior Skyway Aircraft Rapid City, South Dakota

Web Services:

1 The ultimate goal of Web services: Allow different software in different places, written in different languages and resident on different platforms, to connect and interoperate

1 The Web began as provider of markup documents, served through the HTTP methods, GET and POST

1 A Web service is closely related to an information service

- The server provides services, through server- resident software

- The same Web server can provide both documents and services

1 The original Web services were provided via Remote Procedure Call (RPC), through two technologies, DCOM and CORBA. DCOM and CORBA use different protocols, which defeats the goal of universal component interoperability

1 There are three roles required to provide and use Web services:

1. Service providers

2. Service requestors

3. A service registry

- Service providers

- Must develop & deploy software that provide service

- Service Description --> Web Services Definition Language (WSDL)

*Used to describe available services, as well as of message protocols for their use their use

* Such descriptions reside on the Web server

- Service requestors

- Uses WSDL to query a web registry

- Service registry

- Created using Universal Description, Discovery, and Integration Service (UDDI)

* UDDI also provides

- Create service registry
- Provides method to query a Web service registry
- Standard Object Access Protocol (SOAP)
 - An XML-based specification that defines the forms of messages and RPCs
 - The root element of SOAP is envelope
 - Envelope contains SOAP messages – description of web services
 - Supports the exchange of information among distributed systems

UNIT - 6

PERL, CGI PROGRAMMING

6.1 Origins and uses of Perl

Began in the late 1980s as a more powerful replacement for the capabilities of awk (text file processing) and sh (UNIX system administration)

- Now includes sockets for communications and modules for OOP, among other things
- Now the most commonly used language for CGI, in part because of its pattern matching capabilities
- Perl programs are usually processed the same way as many Java programs, compilation to an intermediate form, followed by interpretation

6.2 Scalars and their operations

- Scalars are variables that can store either numbers, strings, or references (discussed later)

- Numbers are stored in double format; integers are rarely used

- Numeric literals have the same form as in other common languages

Perl has two kinds of string literals, those delimited by double quotes and those delimited by single quotes

- Single-quoted literals cannot include escape sequences

- Double-quoted literals can include them

- In both cases, the delimiting quote can be embedded by preceding it with a backslash

- If you want a string literal with single-quote characteristics, but don't want delimit it with single quotes, use qx, where x is a new delimiter

- For double quotes, use qq

- If the new delimiter is a parenthesis, a brace, a bracket, or a pointed bracket, the right delimiter must be the other member of the pair

- A null string can be " or ""

Scalar type is specified by preceding the name with a \$

- Name must begin with a letter; any number of letters, digits, or underscore characters can follow

- Names are case sensitive

- By convention, names of variables use only lowercase letters

- Names embedded in double-quoted string literals are interpolated

e.g., If the value of \$salary is 47500, the value of

"Jack makes \$salary dollars per year" is "Jack makes 47500 dollars per year"

- Variables are implicitly declared

- A scalar variable that has not been assigned a value has the value undef (numeric value is 0; string value is the null string)

- Perl has many implicit variables, the most common

 - of which is \$_ (Look at perldoc perlvar)

- Numeric Operators

- Like those of C, Java, etc.

Operator Associativity

++, -- nonassociative

unary - right

** right

*, /, % left

binary +, - left

- String Operators

- Catenation - denoted by a period e.g., If the value of \$dessert is "apple", the value of \$dessert . " pie" is "apple pie"

- Repetition - denoted by x e.g., If the value of \$greeting is "hello ", the value of

 - \$greeting x 3 is "hello hello hello "

- String Functions

- Functions and operators are closely related in Perl

- e.g., if cube is a predefined function, it can be called with either cube(x) or cube x Name
Parameters Result chomp a string the string w/terminating newline characters removed

length a string the number of characters in the string lc a string the string with uppercase letters converted to lower uc a string the string with lowercase letters converted to upper hex a string the decimal value of the hexadecimal number in the string join a character and the strings catenated a list of strings together with the character inserted between them

6.3 Control statements

In the last chapter you learned how to decode form data, and mail it to yourself. However, one problem with the guestbook program is that it didn't do any error-checking or specialized processing. You might not want to get blank forms, or you may want to require certain fields to be filled out. You might also want to write a quiz or questionnaire, and have your program take different actions depending on the answers. All of these things require some more advanced processing of the form data, and that will usually involve using control structures in your Perl code.

Control structures include conditional statements, such as if/elsif/else blocks, as well as loops like foreach, for and while.

If Conditions

You've already seen if/elsif in action. The structure is always started by the word if, followed by a condition to be evaluated, then a pair of braces indicating the beginning and end of the code to be executed if the condition is true. The condition is enclosed in parentheses:

```
if (condition) {  
    code to be executed  
}
```

The condition statement can be anything that evaluates to true or false. In Perl, any string is true except the empty string and 0. Any number is true except 0. An undefined value

(or `undef`) is false. You can also test whether a certain value equals something, or doesn't equal something, or is greater than or less than something. There are different conditional test operators, depending on whether the variable you want to test is a string or a number:

Relational and Equality Operators

Test	Numbers	Strings
<code>\$x</code> is equal to <code>\$y</code>	<code>\$x == \$y</code>	<code>\$x eq \$y</code>
<code>\$x</code> is not equal to <code>\$y</code>	<code>\$x != \$y</code>	<code>\$x ne \$y</code>
<code>\$x</code> is greater than <code>\$y</code>	<code>\$x > \$y</code>	<code>\$x gt \$y</code>
<code>\$x</code> is greater than or equal to <code>\$y</code>	<code>\$x >= \$y</code>	<code>\$x ge \$y</code>
<code>\$x</code> is less than <code>\$y</code>	<code>\$x < \$y</code>	<code>\$x lt \$y</code>
<code>\$x</code> is less than or equal to <code>\$y</code>	<code>\$x <= \$y</code>	<code>\$x le \$y</code>

If it's a string test, you use the letter operators (`eq`, `ne`, `lt`, etc.), and if it's a numeric test, you use the symbols (`==`, `!=`, etc.). Also, if you are doing numeric tests, keep in mind that `$x >= $y` is not the same as `$x => $y`. Be sure to use the correct operator!

Here is an example of a numeric test. If `$varname` is greater than 23, the code inside the curly braces is executed:

```
if ($varname > 23) {
    # do stuff here if the condition is true
}
```

If you need to have more than one condition, you can add `elsif` and `else` blocks:

```
if ($varname eq "somestring") {
    # do stuff here if the condition is true
}
elsif ($varname eq "someotherstring") {
    # do other stuff
}
```

```

else {
    # do this if none of the other conditions are met
}

```

The line breaks are not required; this example is just as valid:

```

if ($varname > 23) {
    print "$varname is greater than 23";
} elsif ($varname == 23) {
    print "$varname is 23";
} else { print "$varname is less than 23"; }

```

You can join conditions together by using logical operators:

Logical Operators

Operator	Example	Explanation
&&	condition1 condition2	&& True if condition1 and condition2 are both true
	condition1 condition2	True if either condition1 or condition2 is true
and	condition1 and condition2	Same as && but lower precedence
or	condition1 or condition2	Same as but lower precedence

Logical operators are evaluated from left to right. Precedence indicates which operator is evaluated first, in the event that more than one operator appears on one line. In a case like this:

condition1 || condition2 && condition3

condition2 && condition3 is evaluated first, then the result of that evaluation is used in the || evaluation.

and and or work the same way as `&&` and `||`, although they have lower precedence than their symbolic counterparts.

Unless

`unless` is similar to `if`. Let's say you wanted to execute code only if a certain condition were false. You could do something like this:

```
if ($varname != 23) {  
    # code to execute if $varname is not 23  
}
```

The same test can be done using `unless`:

```
unless ($varname == 23) {  
    # code to execute if $varname is not 23  
}
```

There is no "elseunless", but you can use an `else` clause:

```
unless ($varname == 23) {  
    # code to execute if $varname is not 23  
} else {  
    # code to execute if $varname IS 23  
}
```

Validating Form Data

You should always validate data submitted on a form; that is, check to see that the form fields aren't blank, and that the data submitted is in the format you expected. This is typically done with `if/elsif` blocks.

Here are some examples. This condition checks to see if the "name" field isn't blank:

```
if (param('name') eq "") {
```

```

    &dienice("Please fill out the field for your name.");
}


```

You can also test multiple fields at the same time:

```

if (param('name') eq "" or param('email') eq "") {
    &dienice("Please fill out the fields for your name
and email address.");
}


```

The above code will return an error if either the name or email fields are left blank.

`param('fieldname')` always returns one of the following:

undef —	or <code>fieldname</code> is not defined in the form itself, or it's a checkbox/radio button field that wasn't checked.
the empty string	<code>fieldname</code> exists in the form but the user didn't type anything into that field (for text fields)
one or more values	whatever the user typed into the field(s)

If your form has more than one field containing the same `fieldname`, then the values are stored sequentially in an array, accessed by `param('fieldname')`.

You should always validate all form data — even fields that are submitted as hidden fields in your form. Don't assume that your form is always the one calling your program. Any external site can send data to your CGI. Never trust form input data.

Looping

Loops allow you to repeat code for as long as a condition is met. Perl has several loop control structures: `foreach`, `for`, `while` and `until`.

Foreach Loops

foreach iterates through a list of values:

```
foreach my $i (@arrayname) {  
    # code here  
}
```

This loops through each element of @arrayname, setting \$i to the current array element for each pass through the loop. You may omit the loop variable \$i:

```
foreach (@arrayname) {  
    # $_ is the current array element  
}
```

This sets the special Perl variable \$_ to each array element. \$_ does not need to be declared (it's part of the Perl language) and its scope localized to the loop itself.

For Loops

Perl also supports C-style for loops:

```
for ($i = 1; $i < 23; $i++) {  
    # code here  
}
```

The for statement uses a 3-part conditional: the loop initializer; the loop condition (how long to run the loop); and the loop re-initializer (what to do at the end of each iteration of the loop). In the above example, the loop initializes with \$i being set to 1. The loop will run for as long as \$i is less than 23, and at the end of each iteration \$i is incremented by 1 using the auto-increment operator (++).

The conditional expressions are optional. You can do infinite loops by omitting all three conditions:

```
for (;;) {  
    # code here  
}
```

You can also write infinite loops with while.

While Loops

A while loop executes as long as particular condition is true:

```
while (condition) {  
    # code to run as long as condition is true  
}
```

Until Loops

until is the reverse of while. It executes as long as a particular condition is NOT true:

```
until (condition) {  
    # code to run as long as condition is not true  
}
```

Infinite Loops

An infinite loop is usually written like so:

```
while (1) {  
    # code here  
}
```

Obviously unless you want your program to run forever, you'll need some way to break out of these infinite loops. We'll look at breaking next.

Breaking from Loops

There are several ways to break from a loop. To stop the current loop iteration (and move on to the next one), use the next command:

```
foreach my $i (1..20) {  
    if ($i == 13) {  
        next;  
    }  
    print "$i\n";  
}
```

This example prints the numbers from 1 to 20, except for the number 13. When it reaches 13, it skips to the next iteration of the loop.

To break out of a loop entirely, use the last command:

```
foreach my $i (1..20) {  
    if ($i == 13) {  
        last;  
    }  
    print "$i\n";  
}
```

This example prints the numbers from 1 to 12, then terminates the loop when it reaches 13.

next and last only effect the innermost loop structure, so if you have something like this:

```
foreach my $i (@list1) {  
    foreach my $j (@list2) {  
        if ($i == 5 && $j == 23) {  
            last;  
        }  
    }  
}
```

```
# this is where that last sends you
}
```

The last command only terminates the innermost loop. If you want to break out of the outer loop, you need to use loop labels:

```
OUTER: foreach my $i (@list1) {
    INNER: foreach my $j (@list2) {
        if ($i == 5 && $j == 23) {
            last OUTER;
        }
    }
}
# this is where that last sends you
```

The loop label is a string that appears before the loop command (foreach, for, or while). In this example we used OUTER as the label for the outer foreach loop and INNER for the inner loop label.

Now that you've seen the various types of Perl control structures, let's look at how to apply them to handling advanced form data.

6.4 Fundamentals of arrays

An array stores an ordered list of values. While a scalar variable can only store one value, an array can store many. Perl array names are prefixed with an @-sign. Here is an example:

```
my @colors = ("red", "green", "blue");
```

Each individual item (or element) of an array may be referred to by its index number. Array indices start with 0, so to access the first element of the array @colors, you use

`$colors[0]`. Notice that when you're referring to a single element of an array, you prefix the name with `$` instead of `@`. The `$`-sign again indicates that it's a single (scalar) value; the `@`-sign means you're talking about the entire array.

If you want to loop through an array, printing out all of the values, you could print each element one at a time:

```
my @colors = ("red", "green", "blue");

print "$colors[0]\n";          # prints "red"
print "$colors[1]\n";          # prints "green"
print "$colors[2]\n";          # prints "blue"
```

A much easier way to do this is to use a foreach loop:

```
my @colors = ("red", "green", "blue");
foreach my $i (@colors) {
    print "$i\n";
}
```

For each iteration of the foreach loop, `$i` is set to an element of the `@colors` array. In this example, `$i` is "red" the first time through the loop. The braces `{ }` define where the loop begins and ends, so for any code appearing between the braces, `$i` is set to the current loop iterator.

Notice we've used `my` again here to declare the variables. In the foreach loop, `my $i` declares the loop iterator (`$i`) and also limits its scope to the foreach loop itself. After the loop completes, `$i` no longer exists.

We'll cover loops more in Chapter 5.

Getting Data Into And Out Of Arrays

An array is an ordered list of elements. You can think of it like a group of people standing in line waiting to buy tickets. Before the line forms, the array is empty:

```
my @people = ();
```

Then Howard walks up. He's the first person in line. To add him to the @people array, use the push function:

```
push(@people, "Howard");
```

Now Sara, Ken, and Josh get in line. Again they are added to the array using the push function. You can push a list of values onto the array:

```
push(@people, ("Sara", "Ken", "Josh"));
```

This pushes the list containing "Sara", "Ken" and "Josh" onto the end of the @people array, so that @people now looks like this: ("Howard", "Sara", "Ken", "Josh")

Now the ticket office opens, and Howard buys his ticket and leaves the line. To remove the first item from the array, use the shift function:

```
my $who = shift(@people);
```

This sets \$who to "Howard", and also removes "Howard" from the @people array, so @people now looks like this: ("Sara", "Ken", "Josh")

Suppose Josh gets paged, and has to leave. To remove the last item from the array, use the pop function:

```
my $who = pop(@people);
```

This sets \$who to "Josh", and @people is now ("Sara", "Ken")

Both shift and pop change the array itself, by removing an element from the array.

If you want to find out how many elements are in a given array, you can use the scalar function:

```
my @people = ("Howard", "Sara", "Ken", "Josh");
my $linelen = scalar(@people);
print "There are $linelen people in line.\n";
```

This prints "There are 4 people in line." Of course, there's always more than one way to do things in Perl, and that's true here — the scalar function is not actually needed. All you have to do is evaluate the array in a scalar context. You can do this by assigning it to a scalar variable:

```
my $linelen = @people;
```

This sets \$linelen to 4.

What if you want to print the name of the last person in line? Remember that Perl array indices start with 0, so the index of the last element in the array is actually length-1:

```
print "The last person in line is $people[$linelen-1].\n";
```

Perl also has a handy shortcut for finding the index of the last element of an array, the \$# shortcut:

```
print "The last person in line is $people[$#people].\n";
```

\$#arrayname is equivalent to scalar(@arrayname)-1. This is often used in foreach loops where you loop through an array by its index number:

```
my @colors = ("cyan", "magenta", "yellow", "black");
foreach my $i (0..$#colors) {
    print "color $i is $colors[$i]\n";
}
```

This will print out "color 0 is cyan, color 1 is magenta", etc.

The \$#arrayname syntax is one example where an #-sign does not indicate a comment.

Array Slices

You can retrieve part of an array by specifying the range of indices to retrieve:

```
my @colors = ("cyan", "magenta", "yellow", "black");
my @slice = @colors[1..2];
```

This example sets @slice to ("magenta", "yellow").

Finding An Item In An Array

If you want to find out if a particular element exists in an array, you can use the grep function:

```
my @results = grep(/pattern/, @listname);
```

/pattern/ is a regular expression for the pattern you're looking for. It can be a plain string, such as /Box kite/, or a complex regular expression pattern.

/pattern/ will match partial strings inside each array element. To match the entire array element, use /^pattern\$/, which anchors the pattern match to the beginning (^) and end (\$) of the string.

grep returns a list of the elements that matched the pattern.

Sorting Arrays

You can do an alphabetical (ASCII) sort on an array of strings using the sort function:

```
my @colors = ("cyan", "magenta", "yellow", "black");
my @colors2 = sort(@colors);
```

@colors2 becomes the @colors array in alphabetically sorted order ("black", "cyan", "magenta", "yellow"). Note that the sort function, unlike push and pop, does not change

the original array. If you want to save the sorted array, you have to assign it to a variable.

If you want to save it back to the original array variable, you'd do:

```
@colors = sort @colors;
```

You can invert the order of the array with the reverse function:

```
my @colors = ("cyan", "magenta", "yellow", "black");
@colors = reverse(@colors);
```

@colors is now ("black", "yellow", "magenta", "cyan").

To do a reverse sort, use both functions:

```
my @colors = ("cyan", "magenta", "yellow", "black");
@colors = reverse(sort(@colors));
```

@colors is now ("yellow", "magenta", "cyan", "black").

The sort function, by default, compares the ASCII values of the array elements (see <http://www.cgi101.com/book/ch2/ascii.html> for the chart of ASCII values). This means if you try to sort a list of numbers, you get "12" before "2". You can do a true numeric sort like so:

```
my @numberlist = (8, 4, 3, 12, 7, 15, 5);
my @sortednumberlist = sort( { $a <=> $b; } @numberlist);
```

{ \$a <=> \$b; } is actually a small subroutine, embedded right in your code, that gets called for each pair of items in the array. It compares the first number (\$a) to the second number (\$b) and returns a number indicating whether \$a is greater than, equal to, or less than \$b. This is done repeatedly with all the numbers in the array until the array is completely sorted.

Joining Array Elements Into A String

You can merge an array into a single string using the join function:

```
my @colors = ("cyan", "magenta", "yellow", "black");
my $colorstring = join(", ", @colors);
```

This joins @colors into a single string variable (\$colorstring), with each element of the @colors array combined and separated by a comma and a space. In this example \$colorstring becomes "cyan, magenta, yellow, black".

You can use any string (including the empty string) as the separator. The separator is the first argument to the join function:

```
join(separator, list);
```

The opposite of join is split, which splits a string into a list of values. See Chapter 7 for more on split.

Array or List?

In general, any function or syntax that works for arrays will also work for a list of values:

```
my $color = ("red", "green", "blue")[1];
# $color is "green"
```

```
my $colorstring = join(", ", ("red", "green", "blue"));
# $colorstring is now "red, green, blue"
```

```
my ($first, $second, $third) = sort("red", "green", "blue");
# $first is "blue", $second is "green", $third is "red"
```

6.5 Hashes

A hash is a special kind of array — an associative array, or paired list of elements. Each pair consists of a string key and a data value.

Perl hash names are prefixed with a percent sign (%). Here's how they're defined:

Hash Name key value

```
my %colors = ( "red", "#ff0000",
                "green", "#00ff00",
                "blue", "#0000ff",
                "black", "#000000",
                "white", "#ffffff" );
```

This particular example creates a hash named %colors which stores the RGB HEX values for the named colors. The color names are the hash keys; the hex codes are the hash values.

Remember that there's more than one way to do things in Perl, and here's the other way to define the same hash:

Hash Name key value

```
my %colors = ( red => "#ff0000",
                green => "#00ff00",
                blue => "#0000ff",
                black => "#000000",
                white => "#ffffff" );
```

The => operator automatically quotes the left side of the argument, so enclosing quotes around the key names are not needed.

To refer to the individual elements of the hash, you'll do:

```
$colors{'red'}
```

Here, "red" is the key, and \$colors{'red'} is the value associated with that key. In this case, the value is "#ff0000".

You don't usually need the enclosing quotes around the value, either; \$colors{red} also works if the key name doesn't contain characters that are also Perl operators (things like +, -, =, * and /).

To print out all the values in a hash, you can use a foreach loop:

```
foreach my $color (keys %colors) {
    print "$colors{$color}=$color\n";
}
```

This example uses the keys function, which returns a list of the keys of the named hash. One drawback is that keys %hashname will return the keys in unpredictable order — in this example, keys %colors could return ("red", "blue", "green", "black", "white") or ("red", "white", "green", "black", "blue") or any combination thereof. If you want to print out the hash in exact order, you have to specify the keys in the foreach loop:

```
foreach my $color ("red","green","blue","black","white") {
    print "$colors{$color}=$color\n";
}
```

Let's write a CGI program using the colors hash. Start a new file called colors.cgi:

Program 2-2: colors.cgi - Print Hash Variables Program

```
#!/usr/bin/perl -wT
use CGI qw(:standard);
use CGI::Carp qw(warningsToBrowser fatsalsToBrowser);
use strict;

# declare the colors hash:
```

```

my %colors = (      red => "#ff0000", green=> "#00ff00",
                  blue => "#0000ff",           black => "#000000",
                  white => "#ffffff" );

# print the html headers
print header;
print start_html("Colors");

foreach my $color (keys %colors) {
    print "<font color=\"$colors{$color}\">$color</font>\n";
}
print end_html;

```

Save it and chmod 755 colors.cgi, then test it in your web browser.

Notice we've had to add backslashes to escape the quotes in this double-quoted string:

```
print "<font color=\"$colors{$color}\">$color</font>\n";
```

A better way to do this is to use Perl's qq operator:

```
print qq(<font color="$colors{$color}">$color</font>\n);
```

qq creates a double-quoted string for you. And it's much easier to read without all those backslashes in there.

Adding Items to a Hash

To add a new value to a hash, you simply do:

```
$hashname{newkey} = newvalue;
```

Using our colors example again, here's how to add a new value with the key "purple":

```
$colors{purple} = "#ff00ff";
```

If the named key already exists in the hash, then an assignment like this overwrites the previous value associated with that key.

Determining Whether an Item Exists in a Hash

You can use the exists function to see if a particular key/value pair exists in the hash:

```
exists $hashname{key}
```

This returns a true or false value. Here's an example of it in use:

```
if (exists $colors{purple}) {  
    print "Sorry, the color purple is already in the hash.<br>\n";  
} else {  
    $colors{purple} = "#ff00ff";  
}
```

This checks to see if the key "purple" is already in the hash; if not, it adds it.

Deleting Items From a Hash

You can delete an individual key/value pair from a hash with the delete function:

```
delete $hashname{key};
```

If you want to empty out the entire hash, do:

```
%hashname = ();
```

Values

We've already seen that the keys function returns a list of the keys of a given hash.

Similarly, the values function returns a list of the hash values:

```
my %colors = (red => "#ff0000", green=> "#00ff00",  
              blue => "#0000ff", black => "#000000",
```

```

white => "#ffffff" );

my @keyslice = keys %colors;
# @keyslice now equals a randomly ordered list of
# the hash keys:
# ("red", "green", "blue", "black", "white")

my @valueslice = values %colors;
# @valueslice now equals a randomly ordered list of
# the hash values:
# ("ff0000", "#00ff00", "#0000ff", "#000000", "#ffffff")

```

As with keys, values returns the values in unpredictable order.

Determining Whether a Hash is Empty

You can use the scalar function on hashes as well:

```
scalar($hashname);
```

This returns true or false value — true if the hash contains any key/value pairs. The value returned does not indicate how many pairs are in the hash, however. If you want to find that number, use:

```
scalar keys(%hashname);
```

Here's an example:

```

my %colors = (red => "#ff0000", green=> "#00ff00",
              blue => "#0000ff", black => "#000000",
              white => "#ffffff" );

my $numcolors = scalar(keys(%colors));
print "There are $numcolors in this hash.\n";

```

This will print out "There are 5 colors in this hash."

6.7 Functions

The real power of PHP comes from its functions. In PHP, there are more than 700 built-in functions. To keep the script from being executed when the page loads, you can put it into a function. A function will be executed by a call to the function. You may call a function from anywhere within a page.

Create a PHP Function

A function will be executed by a call to the function.

Syntax

```
function functionName()  
{  
    code to be executed;  
}
```

PHP function guidelines:

- Give the function a name that reflects what the function does
- The function name can start with a letter or underscore (not a number)

Example

A simple function that writes my name when it is called:

```
<html>
```

```
<body>
```

```
<?php
```

```
function writeName()  
  
{  
  
echo "Kai Jim Refsnes";  
  
}  
  
echo "My name is ";  
  
writeName();  
?>  
</body>  
</html>
```

Output:

My name is Kai Jim Refsnes

PHP Functions - Adding parameters

To add more functionality to a function, we can add parameters. A parameter is just like a variable. Parameters are specified after the function name, inside the parentheses.

Example 1

The following example will write different first names, but equal last name:

```
<html>  
<body>  
<?php function writeName($fname)  
  
{  
  
echo $fname .
```

```
" Refsnes.<br />";  
}  
  
echo "My name is ";writeName("Kai Jim");  
  
echo "My sister's name is ";  
  
writeName("Hege");  
  
echo "My brother's name is";  
  
writeName("Stale");?></body></html>
```

Output:

My name is Kai Jim Refsnes.

My sister's name is Hege Refsnes.

My brother's name is Stale Refsnes.

Example 2

The following function has two parameters:

```
<html>  
  
<body>  
  
<?php function writeName($fname,$punctuation)  
{  
  
echo $fname . " Refsnes" . $punctuation . "<br />";  
  
}
```

```
echo "My name is ";

writeName("Kai Jim", ".");

echo "My sister's name is ";

writeName("Hege", "!");
echo "My brother's name is ";

writeName("Ståle", "?");

?>

</body>

</html>
```

Output:

My name is Kai Jim Refsnes.

My sister's name is Hege Refsnes!

My brother's name is Ståle Refsnes?

PHP Functions - Return values

To let a function return a value, use the return statement.

Example

```
<html>
<body>

<?php
function add($x,$y)
{
$total=$x+$y;
return $total;
```

```
}
```

```
echo "1 + 16 = " . add(1,16);
?>
```

```
</body>
</html>
```

Output:

1 + 16 = 17

6.8 Pattern matching

Pattern-Matching Operators

Zoologically speaking, Perl's pattern-matching operators function as a kind of cage for regular expressions, to keep them from getting out. This is by design; if we were to let the regex beasties wander throughout the language, Perl would be a total jungle. The world needs its jungles, of course--they're the engines of biological diversity, after all--but jungles should stay where they belong. Similarly, despite being the engines of combinatorial diversity, regular expressions should stay inside pattern match operators where they belong. It's a jungle in there.

As if regular expressions weren't powerful enough, the m// and s/// operators also provide the (likewCSE confined) power of double-quote interpolation. Since patterns are parsed like double-quoted strings, all the normal double-quote conventions will work, including variable interpolation (unless you use single quotes as the delimiter) and special characters indicated with backslash escapes. (See "Specific Characters" later in this chapter.) These are applied before the string is interpreted as a regular expression. (This is one of the few places in the Perl language where a string undergoes more than one pass of processing.) The first pass is not quite normal double-quote interpolation, in that it knows what it should interpolate and what it should pass on to the regular expression

parser. So, for instance, any \$ immediately followed by a vertical bar, closing parenthesis, or the end of the string will be treated not as a variable interpolation, but as the traditional regex assertion meaning end-of-line. So if you say:

```
$foo = "bar";  
/$foo$/;
```

the double-quote interpolation pass knows that those two \$ signs are functioning differently. It does the interpolation of \$foo, then hands this to the regular expression parser:

```
/bar$/;
```

Another consequence of this two-pass parsing is that the ordinary Perl tokener finds the end of the regular expression first, just as if it were looking for the terminating delimiter of an ordinary string. Only after it has found the end of the string (and done any variable interpolation) is the pattern treated as a regular expression. Among other things, this means you can't "hide" the terminating delimiter of a pattern inside a regex construct (such as a character class or a regex comment, which we haven't covered yet). Perl will see the delimiter wherever it is and terminate the pattern at that point.

You should also know that interpolating variables into a pattern slows down the pattern matcher, because it feels it needs to check whether the variable has changed, in case it has to recompile the pattern (which will slow it down even further). See "Variable Interpolation" later in this chapter.

The tr/// transliteration operator does not interpolate variables; it doesn't even use regular expressions! (In fact, it probably doesn't belong in this chapter at all, but we couldn't think of a better place to put it.) It does share one feature with m// and s///, however: it binds to variables using the =~ and !~ operators.

The =~ and !~ operators, described in Chapter 3, "Unary and Binary Operators", bind the scalar expression on their lefthand side to one of three quote-like operators on their right: m// for matching a pattern, s/// for substituting some string for a substring matched by a pattern, and tr/// (or its synonym, y///) for transliterating one set of characters to another

set. (You may write `m//` as `//`, without the `m`, if slashes are used for the delimiter.) If the righthand side of `=~` or `!~` is none of these three, it still counts as a `m//` matching operation, but there'll be no place to put any trailing modifiers (see "Pattern Modifiers" later), and you'll have to handle your own quoting:

```
print "matches" if $somestring =~ $somepattern;
```

Really, there's little reason not to spell it out explicitly:

```
print "matches" if $somestring =~ m/$somepattern/;
```

When used for a matching operation, `=~` and `!~` are sometimes pronounced "matches" and "doesn't match" respectively (although "contains" and "doesn't contain" might cause less confusion).

Apart from the `m//` and `s///` operators, regular expressions show up in two other places in Perl. The first argument to the `split` function is a special match operator specifying what not to return when breaking a string into multiple substrings. See the description and examples for `split` in Chapter 29, "Functions". The `qr//` ("quote regex") operator also specifies a pattern via a regex, but it doesn't try to match anything (unlike `m//`, which does). Instead, the compiled form of the regex is returned for future use. See "Variable Interpolation" for more information.

You apply one of the `m//`, `s///`, or `tr///` operators to a particular string with the `=~` binding operator (which isn't a real operator, just a kind of topicalizer, linguistically speaking).

Here are some examples:

```
$haystack =~ m/needle/           # match a simple pattern
```

```
$haystack =~ /needle/           # same thing
```

```
$italiano =~ s/butter/olive oil/    # a healthy substitution
```

```
$rotate13 =~ tr/a-zA-Z/n-za-mN-ZA-M/ # easy encryption (to break)
```

Without a binding operator, `$_` is implicitly used as the "topic":

```
/new life/ and      # search in $_ and (if found)
```

```
/new civilizations/ # boldly search $_[
```

```
s/sugar/aspartame/ # substitute a substitute into $_[
```

```
tr/ATCG/TAGC/ # complement the DNA stranded in $_[
```

Because s/// and tr/// change the scalar to which they're applied, you may only use them on valid lvalues:

```
"onshore" =~ s/on/off/; # WRONG: compile-time error
```

However, m// works on the result of any scalar expression:

```
if ((lc $magic_hat->fetch_contents->as_string) =~ /rabbit/) {
    print "Nyaa, what's up doc?\n";
}
else {
    print "That trick never works!\n";
}
```

But you have to be a wee bit careful, since =~ and !~ have rather high precedence--in our previous example the parentheses are necessary around the left term.[3] The !~ binding operator works like =~, but negates the logical result of the operation:

```
if ($song !~ /words/) {
    print qq/"$song" appears to be a song without words.\n/;
```

Since m//, s///, and tr/// are quote operators, you may pick your own delimiters. These work in the same way as the quoting operators q//, qq//, qr//, and qw// (see the section Section 5.6.3, "Pick Your Own Quotes" in Chapter 2, "Bits and Pieces").

```
$path =~ s#/tmp#/var/tmp/scratch#;
```

```
if ($dir =~ m[/bin]) {
    print "No binary directories please.\n";
}
```

When using paired delimiters with s/// or tr///, if the first part is one of the four customary bracketing pairs (angle, round, square, or curly), you may choose different delimiters for the second part than you chose for the first:

```
s(egg)<larva>;
```

```
s{larva}{pupa};
```

```
s[pupa]/imago/;
```

Whitespace is allowed in front of the opening delimiters:

```
s (egg) <larva>;
```

```
s {larva} {pupa};
```

```
s [pupa] /imago/;
```

Each time a pattern successfully matches (including the pattern in a substitution), it sets the \$`, \$&, and \$' variables to the text left of the match, the whole match, and the text right of the match. This is useful for pulling apart strings into their components:

```
"hot cross buns" =~ /cross/;
```

```
print "Matched: <$`> $& <$'>\n"; # Matched: <hot> cross <buns>
```

```
print "Left: <$`>\n"; # Left: <hot>
```

```
print "Match: <$&>\n"; # Match: <cross>
```

```
print "Right: <$'>\n"; # Right: <buns>
```

For better granularity and efficiency, use parentheses to capture the particular portions that you want to keep around. Each pair of parentheses captures the substring corresponding to the subpattern in the parentheses. The pairs of parentheses are numbered from left to right by the positions of the left parentheses; the substrings corresponding to those subpatterns are available after the match in the numbered variables, \$1, \$2, \$3, and so on:[4]

```
$_ = "Bilbo Baggins's birthday is September 22";
```

```
/(.*)'s birthday is (.*)/;
```

```
print "Person: $1\n";
```

```
print "Date: $2\n";
```

\$`, \$&, \$', and the numbered variables are global variables implicitly localized to the enclosing dynamic scope. They last until the next successful pattern match or the end of

the current scope, whichever comes first. More on this later, in a different scope.

[3] Without the parentheses, the lower-precedence `lc` would have applied to the whole pattern match instead of just the method call on the magic hat object.

[4] Not `$0`, though, which holds the name of your program.

Once Perl sees that you need one of `$``, `$&`, or `$'` anywhere in the program, it provides them for every pattern match. This will slow down your program a bit. Perl uses a similar mechanism to produce `$1`, `$2`, and so on, so you also pay a price for each pattern that contains capturing parentheses. (See "Clustering" to avoid the cost of capturing while still retaining the grouping behavior.) But if you never use `$`$&`, or `$'`, then patterns without capturing parentheses will not be penalized. So it's usually best to avoid `$``, `$&`, and `$'` if you can, especially in library modules. But if you must use them once (and some algorithms really appreciate their convenience), then use them at will, because you've already paid the price. `$&` is not so costly as the other two in recent versions of Perl.

6.9 File input and output

As you start to program more advanced CGI applications, you'll want to store data so you can use it later. Maybe you have a guestbook program and want to keep a log of the names and email addresses of visitors, or a page counter that must update a counter file, or a program that scans a flat-file database and draws info from it to generate a page. You can do this by reading and writing data files (often called file I/O).

File Permissions

Most web servers run with very limited permissions; this protects the server (and the system it's running on) from malicious attacks by users or web visitors. On Unix systems, the web process runs under its own userid, typically the "web" or "nobody" user. Unfortunately this means the server doesn't have permission to create files in your directory. In order to write to a data file, you must usually make the file (or the directory where the file will be created) world-writable — or at least writable by the web process userid. In Unix a file can be made world-writable using the **chmod** command:

chmod 666 myfile.dat

To set a directory world-writable, you'd do:

chmod 777 directoryname

See Appendix A for a chart of the various chmod permissions.

Unfortunately, if the file is world-writable, it can be written to (or even deleted) by other users on the system. You should be very cautious about creating world-writable files in your web space, and you should never create a world-writable directory there. (An attacker could use this to install their own CGI programs there.) If you must have a world-writable directory, either use /tmp (on Unix), or a directory outside of your web space. For example if your web pages are in /home/you/public_html, set up your writable files and directories in /home/you.

A much better solution is to configure the server to run your programs with your userid. Some examples of this are CGIwrap (platform independent) and suEXEC (for Apache/Unix). Both of these force CGI programs on the web server to run under the program owner's userid and permissions. Obviously if your CGI program is running with your userid, it will be able to create, read and write files in your directory without needing the files to be world-writable.

The Apache web server also allows the webmaster to define what user and group the server runs under. If you have your own domain, ask your webmaster to set up your domain to run under your own userid and group permissions.

Permissions are less of a problem if you only want to read a file. If you set the file permissions so that it is group- and world-readable, your CGI programs can then safely read from that file. Use caution, though; if your program can read the file, so can the webserver, and if the file is in your webspace, someone can type the direct URL and view the contents of the file. Be sure not to put sensitive data in a publicly readable file.

Opening Files

Reading and writing files is done by opening a file and associating it with a filehandle.

This is done with the statement:

```
open(filehandle,filename);
```

The filename may be prefixed with a `>`, which means to overwrite anything that's in the file now, or with a `>>`, which means to append to the bottom of the existing file. If both `>` and `>>` are omitted, the file is opened for reading only. Here are some examples:

```
open(INF,"out.txt");      # opens mydata.txt for reading
open(OUTF,>"out.txt");    # opens out.txt for overwriting
open(OUTF,>>"out.txt");  # opens out.txt for appending
open(FH, "+<out.txt");   # opens existing file out.txt for reading AND writing
```

The filehandles in these cases are `INF`, `OUTF` and `FH`. You can use just about any name for the filehandle.

Also, a warning: your web server might do strange things with the path your programs run under, so it's possible you'll have to use the full path to the file (such as `/home/you/public_html/somedata.txt`), rather than just the filename. This is generally not the case with the Apache web server, but some other servers behave differently. Try opening files with just the filename first (provided the file is in the same directory as your CGI program), and if it doesn't work, then use the full path.

One problem with the above code is that it doesn't check the return value of `open` to ensure the file was really opened. `open` returns nonzero upon success, or `undef` (which is a false value) otherwise. The safe way to open a file is as follows:

```
open(OUTF,>"outdata.txt") or &dienice("Can't open outdata.txt for writing: $!");
```

This uses the "dienice" subroutine we wrote in Chapter 4 to display an error message and

exit if the file can't be opened. You should do this for all file opens, because if you don't, your CGI program will continue running even if the file isn't open, and you could end up losing data. It can be quite frustrating to realize you've had a survey running for several weeks while no data was being saved to the output file.

The \$! in the above example is a special Perl variable that stores the error code returned by the failed open statement. Printing it may help you figure out why the open failed.

Guestbook Form with File Write

Let's try this by modifying the guestbook program you wrote in Chapter 4. The program already sends you e-mail with the information; we're going to have it write its data to a file as well.

First you'll need to create the output file and make it writable, because your CGI program probably can't create new files in your directory. If you're using Unix, log into the Unix shell, **cd** to the directory where your guestbook program is located, and type the following:

```
touch guestbook.txt  
chmod 622 guestbook.txt
```

The Unix **touch** command, in this case, creates a new, empty file called "guestbook.txt". (If the file already exists, touch simply updates the last-modified timestamp of the file.) The chmod 622 command makes the file read/write for you (the owner), and write-only for everyone else.

If you don't have Unix shell access (or you aren't using a Unix system), you should create or upload an empty file called guestbook.txt in the directory where your guestbook.cgi program is located, then adjust the file permissions on it using your FTP program.

Now you'll need to modify guestbook.cgi to write to the file:

Program 6-1: guestbook.cgi - Guestbook Program With File Write

```
#!/usr/bin/perl -wT
use CGI qw(:standard);
use CGI::Carp qw(warningsToBrowser fatsalsToBrowser);
use strict;

print header;
print start_html("Results");

# first print the mail message...

$ENV{PATH} = "/usr/sbin";
open (MAIL, "/usr/sbin/sendmail -oi -t -odq") or
    &dienice("Can't fork for sendmail: $!\n");
print MAIL "To: recipient@cgi101.com\n";
print MAIL "From: nobody@cgi101.com\n";
print MAIL "Subject: Form Data\n\n";
foreach my $p (param()) {
    print MAIL "$p = ", param($p), "\n";
}
close(MAIL);

# now write (append) to the file

open(OUT, ">>guestbook.txt") or &dienice("Couldn't open output file: $!");
foreach my $p (param()) {
    print OUT param($p), "|";
}
print OUT "\n";
close(OUT);
```

```
print <<EndHTML;
<h2>Thank You</h2>
<p>Thank you for writing!</p>
<p>Return to our <a href="index.html">home page</a>.</p>
EndHTML
```

```
print end_html;
```

```
sub dienice {
    my($errmsg) = @_;
    print "<h2>Error</h2>\n";
    print "<p>$errmsg</p>\n";
    print end_html;
    exit;
}
```

Now go back to your browser and fill out the guestbook form again. If your CGI program runs without any errors, you should see data added to the guestbook.txt file. The resulting file will show the submitted form data in pipe-separated form:

Someone|someone@wherever.com|comments here

Ideally you'll have one line of data (or record) for each form that is filled out. This is what's called a flat-file database.

Unfortunately if the visitor enters multiple lines in the comments field, you'll end up with multiple lines in the data file. To remove the newlines, you should substitute newline characters (\n) as well as hard returns (\r). Perl has powerful pattern matching and replacement capabilities; it can match the most complex patterns in a string using regular expressions (see Chapter 13). The basic syntax for substitution is:

```
$mystring =~ s/pattern/replacement/;
```

This command substitutes "pattern" for "replacement" in the scalar variable \$mystring. Notice the operator is a `=~` (an equals sign followed by a tilde); this is Perl's binding operator and indicates a regular expression pattern match/substitution/replacement is about to follow.

Here is how to replace the end-of-line characters in your guestbook program:

```
foreach my $p (param()) {
    my $value = param($p);
    $value =~ s/\n/g;    # replace newlines with spaces
    $value =~ s/\r/g;    # remove hard returns
    print OUT "$p = $value,";
}
```

Go ahead and change your program, then test it again in your browser. View the guestbook.txt file in your browser or in a text editor and observe the results.

File Locking

CGI processes on a Unix web server can run simultaneously, and if two programs try to open and write the same file at the same time, the file may be erased, and you'll lose all of your data. To prevent this, you need to lock the files you are writing to. There are two types of file locks:

- A shared lock allows more than one program (or other process) to access the file at the same time. A program should use a shared lock when reading from a file.
- An exclusive lock allows only one program or process to access the file while the lock is held. A program should use an exclusive lock when writing to a file.

File locking is accomplished in Perl using the Fcntl module (which is part of the standard library), and the flock function. The use statement is like CGI.pm's:

```
use Fcntl qw(:flock);
```

The Fcntl module provides symbolic values (like abbreviations) representing the correct lock numbers for the flock function, but you must specify: flock in the use statement in order for Fcntl to export those values. The values are as follows:

LOCK_SH	shared lock
LOCK_EX	exclusive lock
LOCK_NB	non-blocking lock
LOCK_UN	unlock

These abbreviations can then be passed to flock. The flock function takes two arguments: the filehandle and the lock type, which is typically a number. The number may vary depending on what operating system you are using, so it's best to use the symbolic values provided by Fcntl. A file is locked after you open it (because the filehandle doesn't exist before you open the file):

```
open(FH, "filename") or &die("Can't open file: $!");  
flock(FH, LOCK_SH);
```

The lock will be released automatically when you close the file or when the program finishes.

Keep in mind that file locking is only effective if all of the programs that read and write to that file also use flock. Programs that don't will ignore the locks held by other processes.

Since flock may force your CGI program to wait for another process to finish writing to a file, you should also reset the file pointer, using the seek function:

```
seek(filehandle, offset, whence);
```

offset is the number of bytes to move the pointer, relative to whence, which is one of the following:

- 0 beginning of file
- 1 current file position
- 2 end of file

So seek(OUTF,0,2) repositions the pointer to the end of the file. If you were reading the file instead of writing to it, you'd want to do seek(OUTF,0,0) to reset the pointer to the beginning of the file.

The Fcntl module also provides symbolic values for the seek pointers:

- SEEK_SET beginning of file
- SEEK_CUR current file position
- SEEK_END end of file

To use these, add :seek to the use Fcntl statement:

```
use Fcntl qw(:flock :seek);
```

Now you can use seek(OUTF,0,SEEK_END) to reset the file pointer to the end of the file, or seek(OUTF,0,SEEK_SET) to reset it to the beginning of the file.

Closing Files

When you're finished writing to a file, it's best to close the file, like so:

```
close(filehandle);
```

Files are automatically closed when your program ends. File locks are released when the file is closed, so it is not necessary to actually unlock the file before closing it. (In fact, releasing the lock before the file is closed can be dangerous and cause you to lose data.)

Reading Files

There are two ways you can handle reading data from a file: you can either read one line at a time, or read the entire file into an array. Here's an example:

```
open(FH,"guestbook.txt") or &dienice("Can't open guestbook.txt: $!");  
  
my $a = <FH>;    # reads one line from the file into  
                  # the scalar $a  
my @b = <FH>;   # reads the ENTIRE FILE into array @b  
  
close(FH);      # closes the file
```

If you were to use this code in your program, you'd end up with the first line of guestbook.txt being stored in \$a, and the remainder of the file in array @b (with each element of @b containing one line of data from the file). The actual read occurs with <filehandle>; the amount of data read depends on the type of variable you save it into.

The following section of code shows how to read the entire file into an array, then loop through each element of the array to print out each line:

```
open(FH,"guestbook.txt") or &dienice("Can't open guestbook.txt: $!");  
my @ary = <FH>;  
close(FH);  
  
foreach my $line (@ary) {  
    print $line;  
}
```

This code minimizes the amount of time the file is actually open. The drawback is it causes your CGI program to consume as much memory as the size of the file. Obviously for very large files that's not a good idea; if your program consumes more memory than the machine has available, it could crash the whole machine (or at the very least make

things extremely slow). To process data from a very large file, it's better to use a while loop to read one line at a time:

```
open(FH,"guestbook.txt") or &dienice("Can't open guestbook.txt: $!");  
while (my $line = <FH>) {  
    print $line;  
}  
close(FH);
```

Poll Program

Let's try another example: a web poll. You've probably seen them on various news sites. A basic poll consists of one question and several potential answers (as radio buttons); you pick one of the answers, vote, then see the poll results on the next page.

Start by creating the poll HTML form. Use whatever question and answer set you wish.

Program 6-2: poll.html - Poll HTML Form

```
<form action="poll.cgi" method="POST">  
Which was your favorite <i>Lord of the Rings</i> film?<br>  
<input type="radio" name="pick" value="fotr">The Fellowship of the Ring<br>  
<input type="radio" name="pick" value="ttt">The Two Towers<br>  
<input type="radio" name="pick" value="rotk">Return of the King<br>  
<input type="radio" name="pick" value="none">I didn't watch them<br>  
<input type="submit" value="Vote">  
</form>  
<a href="results.cgi">View Results</a><br>
```

In this example we're using abbreviations for the radio button values. Our CGI program will translate the abbreviations appropriately.

Now the voting CGI program will write the result to a file. Rather than having this program analyze the results, we'll simply use a redirect to bounce the viewer to a third program (results.cgi). That way you won't need to write the results code twice.

Here is how the voting program (poll.cgi) should look:

Program 6-3: poll.cgi - Poll Program

```
#!/usr/bin/perl -wT
use CGI qw(:standard);
use CGI::Carp qw(warningsToBrowser fatsalsToBrowser);
use strict;
use Fcntl qw(:flock :seek);

my $outfile = "poll.out";

# only record the vote if they actually picked something
if (param('pick')) {
    open(OUT, ">>$outfile") or &dienice("Couldn't open $outfile: $!");
    flock(OUT, LOCK_EX);    # set an exclusive lock
    seek(OUT, 0, SEEK_END); # then seek the end of file
    print OUT param('pick'),"\n";
    close(OUT);
} else {
    # this is optional, but if they didn't vote, you might
    # want to tell them about it...
    &dienice("You didn't pick anything!");
}

# redirect to the results.cgi.
# (Change to your own URL...)
```

```
print redirect("http://cgi101.com/book/ch6/results.cgi");

sub dience {
    my($msg) = @_;
    print header;
    print start_html("Error");
    print h2("Error");
    print $msg;
    print end_html;
    exit;
}
```

Finally results.cgi reads the file where the votes are stored, totals the overall votes as well as the votes for each choice, and displays them in table format.

Program 6-4: results.cgi - Poll Results Program

```
#!/usr/bin/perl -wT
use CGI qw(:standard);
use CGI::Carp qw(warningsToBrowser fatsalsToBrowser);
use strict;
use Fcntl qw(:flock :seek);

my $outfile = "poll.out";

print header;
print start_html("Results");

# open the file for reading
open(IN, "$outfile") or &dienice("Couldn't open $outfile: $!");
# set a shared lock
flock(IN, LOCK_SH);
# then seek the beginning of the file
```

```
seek(IN, 0, SEEK_SET);

# declare the totals variables
my($total_votes, %results);
# initialize all of the counts to zero:
foreach my $i ("fotr", "ttt", "rotk", "none") {
    $results{$i} = 0;
}

# now read the file one line at a time:
while (my $rec = <IN>) {
    chomp($rec);
    $total_votes = $total_votes + 1;
    $results{$rec} = $results{$rec} + 1;
}
close(IN);

# now display a summary:
print <<End;
<b>Which was your favorite <i>Lord of the Rings</i> film?
</b><br>
<table border=0 width=50%>
<tr>
    <td>The Fellowship of the Ring</td>
    <td>$results{fotr} votes</td>
</tr>
<tr>
    <td>The Two Towers</td>
    <td>$results{ttt} votes</td>
</tr>
<tr>
```

```
<td>Return of the King</td>
<td>$results{rotk} votes</td>
</tr>
<tr>
<td>didn't watch them</td>
<td>$results{none} votes</td>
</tr>
</table>
<p>
$total_votes votes total
</p>
End

print end_html;

sub dienice {
    my($msg) = @_;
    print h2("Error");
    print $msg;
    print end_html;
    exit;
}
```

6.10 The Common Gateway Interface

The **Common Gateway Interface (CGI)** is a standard (see RFC3875: CGI Version 1.1) that defines how webserver software can delegate the generation of webpages to a console application. Such applications are known as CGI scripts; they can be written in any programming language, although scripting languages are often used. In simple words the CGI provides an interface between the webservers and the clients.

Purpose

The task of a webserver is to respond to requests for webpages issued by clients (usually web browsers) by analyzing the content of the request (which is mostly in its URL), determining an appropriate document to send in response, and returning it to the client.

If the request identifies a file on disk, the server can just return the file's contents. Alternatively, the document's content can be composed on the fly. One way of doing this is to let a console application compute the document's contents, and tell the web server to use that console application. CGI specifies which information is communicated between the webserver and such a console application, and how.

The webserver software will invoke the console application as a command. CGI defines how information about the request (such as the URL) is passed to the command in the form of arguments and environment variables. The application is supposed to write the output document to standard output; CGI defines how it can pass back extra information about the output (such as the MIME type, which defines the type of document being returned) by prepending it with headers.

6.11 CGI linkage

CGI programs often are stored in a directory named cgi-bin

- Some CGI programs are in machine code, but Perl programs are usually kept in source form, so perl must be run on them
- A source file can be made to be **executable** by adding a line at their beginning that specifies that a language processing program be run on them first

For Perl programs, if the perl system is stored in

/usr/local/bin/perl, as is often is in UNIX

systems, this is

```
#!/usr/local/bin/perl -w
```

- An HTML document specifies a CGI program with the hypertext reference attribute, href, of an anchor tag, <a>, as in

```
<a href =
"http://www.cs.uccs.edu/cgi-bin/reply.pl">
Click here to run the CGI program, reply.pl
</a>
<!-- reply.html - calls a trivial cgi program
-->
<html>
<head>
<title>
HTML to call the CGI-Perl program reply.pl
</title>
</head>
<body>
This is our first CGI-Perl example
<a href =
"http://www.cs.ucp.edu/cgi-bin/reply.pl">
Click here to run the CGI program, reply.pl
</a>
</body>
</html>
```

- The connection from a CGI program back to the requesting browser is through standard output, usually through the server
- The HTTP header needs only the content type, followed by a blank line, as is created with:

```
print "Content-type: text/html \n\n";
#!/usr/local/bin/perl
# reply.pl – a CGI program that returns a
# greeting to the user
print "Content-type: text/html \n\n",
"<html> <head> \n",
"<title> reply.pl example </title>",


```

```
" </head> \n", "<body> \n",
"<h1> Greetings from your Web server!",
" </h1> \n </body> </html> \n";
```

6.12 Query string format

In World Wide Web, a **query string** is the part of a Uniform Resource Locator (URL) that contains data to be passed to web applications such as CGI programs.

The Mozilla URL location bar showing an URL with the query string title=Main_page&action=raw

When a web page is requested via the Hypertext Transfer Protocol, the server locates a file in its file system based on the requested URL. This file may be a regular file or a program. In the second case, the server may (depending on its configuration) run the program, sending its output as the required page. The query string is a part of the URL which is passed to the program. Its use permits data to be passed from the HTTP client (often a web browser) to the program which generates the web page.

Structure

A typical URL containing a query string is as follows:

`http://server/path/program?query_string`

When a server receives a request for such a page, it runs a program (if configured to do so), passing the `query_string` unchanged to the program. The question mark is used as a separator and is not part of the query string.

A link in a web page may have a URL that contains a query string. However, the main use of query strings is to contain the content of an HTML form, also known as web form. In particular, when a form containing the fields `field1`, `field2`, `field3` is submitted, the

content of the fields is encoded as a query string as follows:

field₁=value₁&field₂=value₂&field₃=value₃...

- The query string is composed of a series of field-value pairs.
- The field-value pairs are each separated by an equal sign.
- The series of pairs is separated by the ampersand, '=' or semicolon, ';'.

For each field of the form, the query string contains a pair field=value. Web forms may include fields that are not visible to the user; these fields are included in the query string when the form is submitted

This convention is a W3C recommendation. W3C recommends that all web servers support semicolon separators in the place of ampersand separators.

Technically, the form content is only encoded as a query string when the form submission method is GET. The same encoding is used by default when the submission method is POST, but the result is not sent as a query string, that is, is not added to the action URL of the form. Rather, the string is sent as the body of the request.

URL encoding

Main article: URL encoding

Some characters cannot be part of a URL (for example, the space) and some other characters have a special meaning in a URL: for example, the character '#' can be used to further specify a subsection (or fragment) of a document; the character '=' is used to separate a name from a value. A query string may need to be converted to satisfy these constraints. This can be done using a schema known as URL encoding.

In particular, encoding the query string uses the following rules:

- Letters (A-Z and a-z), numbers (0-9) and the characters '.', '-', '~' and '_' are left as-is
- SPACE is encoded as '+'

- All other characters are encoded as %FF hex representation with any non-ASCII characters first encoded as UTF-8 (or other specified encoding)

The encoding of SPACE as '+' and the selection of "as-is" characters distinguishes this encoding from RFC 1738.

Example

If a form is embedded in an HTML page as follows:

```
<form action="cgi-bin/test.cgi" method="get">
  <input type="text" name="first">
  <input type="text" name="second">
  <input type="submit">
</form>
```

and the user inserts the strings →this is a field← and →was it clear (already)?← in the two text fields and presses the submit button, the program test.cgi will receive the following query string:

first=this+is+a+field&second=was+it+clear+%28already%29%3F

If the form is processed on the server by a CGI script, the script may typically receive the query string as an environment variable named QUERY_STRING.

6.13 CGI.pm module

CGI.pm is a large and widely used Perl module for programming Common Gateway Interface (CGI) web applications, providing a consistent API for receiving user input and producing HTML or XHTML output. The module is written and maintained by Lincoln D. Stein.

A Sample CGI Page

Here is a simple CGI page, written in Perl using CGI.pm (in object oriented style):

```
#!/usr/bin/perl -w
#
use strict;
use warnings;
use CGI;

my $cgi = CGI->new();

print $cgi->header('text/html');
print $cgi->start_html('A Simple CGI Page'),
$cgi->h1('A Simple CGI Page'),
$cgi->start_form,
'Name: ',
$cgi->textfield('name'), $cgi->br,
'Age: ',
$cgi->textfield('age'), $cgi->p,
$cgi->submit('Submit!'),
$cgi->end_form, $cgi->p,
$cgi->hr;

if ( $cgi->param('name') ) {
    print 'Your name is ', $cgi->param('name'), $cgi->br;
}

if ( $cgi->param('age') ) {
    print 'You are ', $cgi->param('age'), ' years old.';
}
```

```
print $cgi->end_html;
```

This would print a very simple webform, asking for your name and age, and after having been submitted, redisplaying the form with the name and age displayed below it. This sample makes use of CGI.pm's object-oriented abilities; it can also be done by calling functions directly, without the \$cgi->.

Note: in many examples \$q, short for query, is used to store a CGI object. As the above example illustrates, this might be very misleading.

Here is another script that produces the same output using CGI.pm's procedural interface:

```
#!/usr/bin/perl
use strict;
use warnings;
use CGI ':standard';

print header,
      start_html('A Simple CGI Page'),
      h1('A Simple CGI Page'),
      start_form,
      'Name: ',
      textfield('name'), br,
      'Age: ',
      textfield('age'), p,
      submit('Submit!'),
      end_form, p,
      hr;

print 'Your name is ', param('name'), br if param 'name';
print 'You are ', param('age'), ' years old.' if param 'age';
```

```
print end_html;
```

6.14 Cookies

Cookie, also known as a **web cookie**, **browser cookie**, and **HTTP cookie**, is a text string stored by a user's web browser. A cookie consists of one or more name-value pairs containing bits of information, which may be encrypted for information privacy and data security purposes.

The cookie is sent as an HTTP header by a web server to a web browser and then sent back unchanged by the browser each time it accesses that server. A cookie can be used for authentication, session tracking (state maintenance), storing site preferences, shopping cart contents, the identifier for a server-based session, or anything else that can be accomplished through storing textual data.

As text, cookies are not executable. Because they are not executed, they cannot replicate themselves and are not viruses. However, due to the browser mechanism to set and read cookies, they can be used as spyware. Anti-spyware products may warn users about some cookies because cookies can be used to track people—a privacy concern.

Most modern browsers allow users to decide whether to accept cookies, and the time frame to keep them, but rejecting cookies makes some websites unusable.

Uses

Session management

Cookies may be used to maintain data related to the user during navigation, possibly across multiple visits. Cookies were introduced to provide a way to implement a "shopping cart" (or "shopping basket"),^{[2][3]} a virtual device into which users can store items they want to purchase as they navigate throughout the site.

Shopping basket applications today usually store the list of basket contents in a database on the server side, rather than storing basket items in the cookie itself. A web server typically sends a cookie containing a unique session identifier. The web browser will send back that session identifier with each subsequent request and shopping basket items are stored associated with a unique session identifier.

Allowing users to log in to a website is a frequent use of cookies. Typically the web server will first send a cookie containing a unique session identifier. Users then submit their credentials and the web application authenticates the session and allows the user access to services.

Personalization

Cookies may be used to remember the information about the user who has visited a website in order to show relevant content in the future. For example a web server may send a cookie containing the username last used to log in to a web site so that it may be filled in for future visits.

Many websites use cookies for personalization based on users' preferences. Users select their preferences by entering them in a web form and submitting the form to the server. The server encodes the preferences in a cookie and sends the cookie back to the browser. This way, every time the user accesses a page, the server is also sent the cookie where the preferences are stored, and can personalize the page according to the user preferences. For example, the Wikipedia website allows authenticated users to choose the webpage skin they like best; the Google search engine allows users (even non-registered ones) to decide how many search results per page they want to see.

Tracking

Tracking cookies may be used to track internet users' web browsing habits. This can also be done in part by using the IP address of the computer requesting the page or the referrer field of the HTTP header, but cookies allow for a greater precision. This can be done for

example as follows:

1. If the user requests a page of the site, but the request contains no cookie, the server presumes that this is the first page visited by the user; the server creates a random string and sends it as a cookie back to the browser together with the requested page;
2. From this point on, the cookie will be automatically sent by the browser to the server every time a new page from the site is requested; the server sends the page as usual, but also stores the URL of the requested page, the date/time of the request, and the cookie in a log file.

By looking at the log file, it is then possible to find out which pages the user has visited and in what sequence. For example, if the log contains some requests done using the cookie id=abc, it can be determined that these requests all come from the same user. The URL and date/time stored with the cookie allows for finding out which pages the user has visited, and at what time.

Third-party cookies and Web bugs, explained below, also allow for tracking across multiple sites. Tracking within a site is typically used to produce usage statistics, while tracking across sites is typically used by advertising companies to produce anonymous user profiles (which are then used to determine what advertCSEments should be shown to the user).

A tracking cookie may potentially infringe upon the user's privacy but they can be easily removed. Current versions of popular web browsers include options to delete 'persistent' cookies when the application is closed.

Third-party cookies

When viewing a Web page, images or other objects contained within this page may reside on servers besides just the URL shown in your browser. While rendering the page, the browser downloads all these objects. Most modern websites that you view contain information from lots of different sources. For example, if you type www.domain.com into your browser, widgets and advertCSEments within this page are often served from a different domain source. While this information is being retrieved, some of these sources

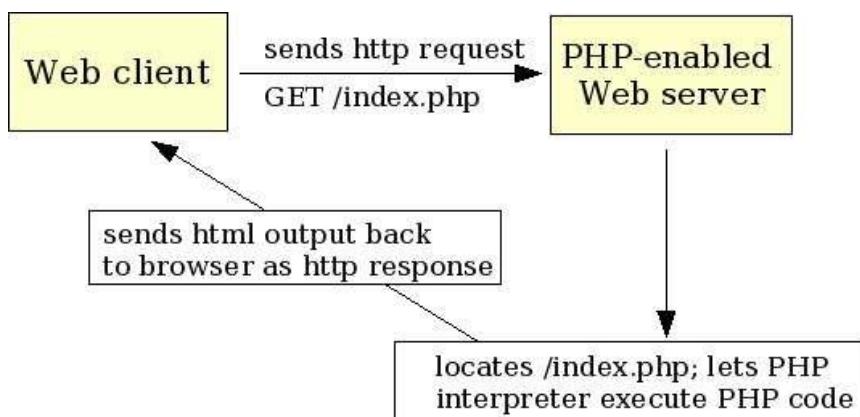
may set cookies in your browser. First-party cookies are cookies that are set by the same domain that is in your browser's address bar. Third-party cookies are cookies being set by one of these widgets or other inserts coming from a different domain.

Modern browsers, such as Mozilla Firefox, Internet Explorer and Opera, by default, allow third-party cookies, although users can change the settings to block them. There is no inherent security risk of third-party cookies (they do not harm the user's computer) and they make lots of functionality of the web possible, however some internet users disable them because they can be used to track a user browsing from one website to another. This tracking is most often done by on-line advertising companies to assist in targeting advertCSEments. For example: Suppose a user visits www.domain1.com and an advertCSEr sets a cookie in the user's browser, and then the user later visits www.domain2.com. If the same company advertCSEs on both sites, the advertCSEr knows that this particular user who is now viewing www.domain2.com also viewed www.domain1.com in the past and may avoid repeating advertCSEments. The advertCSEr does not know anything more about the user than that—they do not know the user's name or address or any other personal information (unless they obtain it from another source such as from the user or by reading another cookie).

UNIT-7

PHP

A PHP program is a text file containing zero or more *statements* (i.e., instructions) that tell the PHP interpreter what to do. When configured properly, the web server on which the program resides hands off certain files – most typically, files named *.php – to the PHP interpreter for execution of any PHP code contained in the file. Then the server sends the resulting *output* (rather than the PHP source code) back to the browser.



One consequence of the fact that PHP pages are dynamically generated is that if you want to develop PHP on your own system, you have to run a web server with PHP support. You test a PHP page by pointing your browser at it on your own server via http; you cannot view the output by accessing it directly through the file system (e.g., file:///C:/some_folder/some_file.php). Fortunately, running your own server is not at all difficult; more about that later.

You can have HTML and PHP code in the same physical file. Opening and closing tags tell the PHP parser where the PHP code begins and ends. Everything else is sent to the browser as is. Just as you can have PHP and HTML together in the same page, PHP will also gladly execute a file consisting of just HTML or just PHP code.

```

<p>Some plain old html</p>
<p>
<?php
    // outputs Hi everybody to the standard output -- usually a browser
    echo "Hi everybody.";
    // assigns the sum of 2 and 2 to the variable $sum
    $sum = 2 + 2;
    // you can probably guess the output of:
    echo " 2 + 2 is $sum";

?>
</p>
<p>More plain old html</p>
  
```

Short tags <? echo "foo"; ?> and asp-style tags <% echo "foo" %> can be turned on a setting in PHP's master configuration file, which by default is called `php.ini`. <?php echo \$foo ?> and <script language="PHP">/></script> are always available; first is most common, and preferred.

Each *instruction* is delimited with a semicolon. The closing ?> tag is equivalent to a semicolon, and therefore optional. in this situation:

```
<?php
    echo "This is a test";
?>

<?php echo "This is a test" ?>
```

Whitespace is not significant; it is used to aid legibility.

Comments

Comments are exactly that. The PHP interpreter ignores them; they are for the benefit of human readers of the source code. They are typically used as a sort narrative to explain what is happening in your code, or like post-it notes to remind you of things to think about during development. Comments are also handy for temporarily disabling lines of code.

Unlike HTML comments, your PHP comments do not become part of the output visible to the browser.

So-called C-style and C++ style are supported:

```
<?php
    echo "This is a test"; // This is a one-line c++ style comment
    /* This is a multi line comment
       yet another line of comment */
    echo "This is yet another test";
?>
...as are shell- or Perl-style comments, with a leading # :
```

```
<?php
# here's a comment
$foo = 23; # another comment here
?>
```

Variables

The notion of variables is fundamental to computer programming. They allow programs to behave dynamically and produce different output based on – well, variable – information that is decided at *runtime* (i.e., when the program is executed as opposed to when it is written (and possibly, compiled)). Variables are often supplied through user input. (Think of Google's search results) They can be compared to boxes with labels on them. The labels are the variable names. The stuff the boxes contain is data of some kind. We typically save stuff in these labeled boxes now so we can identify them and do something with their contents later.

A valid variable name starts with a letter or underscore, followed by any number of letters, numbers, or underscores. You assign a value to a variable with the assignment operator: = Hence:

is how you would store the value 7 in a variable called \$number. We will discuss PHP operators later on.

Data types

PHP has eight "primitive" data types – primitive as in fundamental: anything fancier (more complex) is made out of these.

In general you do not have to be as concerned about the data types of your variables as you do with more strictly typed languages. PHP converts things for you automatically much of the time. There are nonetheless situations where you will need to be aware of the kind of data you are dealing with. (Also, when you work with databases such as MySQL you will find that they are pickier about what types of data you are handing them, so it is useful to start thinking in these terms).

There are two compound data types: **array** and **object**. We will talk about arrays next week and objects in a few weeks.

There are also two special types: *resource* and *NULL*. Resources typically represent a link to something external to PHP, like a database connection or filehandle; we will get involved with these later. The special NULL value represents that a variable has no value at all. NULL is the only possible value of type NULL. It is useful in certain circumstances – we'll see examples in the weeks to come.

PHP has four *scalar* types: **boolean**; **integer**; **float** (aka double); and **string**. Scalar means a single thing, as opposed to a compound or complex thing.

boolean is the simplest type: it has only two possible values, either TRUE or FALSE. You often use some kind of expression that evaluates to a boolean which determines the flow of execution:

```
<?php if ($temperature > 75) { echo "it's too hot in here !" ; } ?>
```

integers are whole numbers (negative or positive). They can be notated in decimal, octal or hexadecimal notation. We'll be using decimal exclusively (or very nearly) so in this class. But you should be aware that in octal notation you precede the number with a zero, like so:

```
<?php  
    $number = 03724;  
    echo $number; // outputs: 2004  
?>
```

And in hexadecimal, 2004 looks like this:

7d4

The max value depends on your operating system; for most applications you will find it is big enough (around 2 billion).

floating point numbers (or simply **floats**) are numbers like with a decimal point, like 1.432. The max value is about 1.8e308

a **string** is a series of characters. This is the data type you will probably work with most frequently, since a great many web pages consist mostly of text. The most common way to tell PHP something is a string is to put it in quotes – which makes sense, if you think about it. Single quotes and double quotes are both supported, but they can produce different results, so the choice is not just aesthetic.

Here, the contents of \$string1 and \$string2 will be identical.

```
<?php
$string1 = "Hello world";
$string2 = 'Hello world';
?>
```

However, the double quotes support *variable interpolation* like so:

```
<?php
$string = 'world';
echo "Hello $string"; // prints "Hello world" because $string gets expanded
// versus //
echo 'Hello $string'; // literally prints Hello $string
?>
```

Further, there are special characters like \n and \t that expand to newline and tab respectively in double-(but not single-) quoted strings. See <http://php.net/manual/en/language.types.string.php> for a complete list.

When you need to embed a literal dollar sign within a double-quoted string, you have to *escape* it by preceding it with a backslash.

```
<?php echo "The value of \$string is now $string"; ?>
```

The same is true when you need to embed a single quote in a single-quoted string, or a double-quote in a double-quoted string.

```
<?php echo 'John\'s cat\'s name is George'; ?>
```

The quotes, whether double or single, have to come in matching pairs, i.e., the first to signify the beginning and the second to signify the end of the string; otherwise you get a parse error. (The same is true of other punctuation characters that PHP has pressed into service: brackets, curly braces, and parentheses -- all of which we will encounter shortly.)

Remember our earlier comments about data types, and octal integer notation? Sometimes you do need to care about what kind of data you are working with:

```
<?php
```

```
?>
```

```
echo "zip is $zip\n";
$zip =           // versus
07302; $zip = '07302';
echo "zip is $zip\n";
```

Above, the quotes tell PHP 'this is a string.' Leave them out, and PHP thinks you mean an integer in octal notation.

You can -- and sometimes you have to -- use curly braces around string-embedded variables to "protect" them. For example:

```
<?php
    $beers = 45;
    $beer  = 'Sierra Nevada Pale Ale';
    echo "it would be excessive to drink $beers {$beer}s";
?>
```

Besides double and single quotes, there is a third way of writing string expressions known as the *heredoc syntax*. Provide an identifier after <<<, then the string, and then the same identifier to close the quotation. Heredoc text behaves just like a double-quoted string, without the double-quotes; variables are expanded. You do not have to escape quotes in your here docs, but you can still use escape codes like \n and \t.

```
<?php
$str = <<<EOD
Example of string spanning multiple lines
using heredoc syntax. You could have $foo
here and it would expand.
EOD;
?>
```

The closing identifier must begin in the first column (the beginning) of the line. Also, the identifier used must follow the same naming rules as any other label in PHP: it must contain only alphanumeric characters and underscores, and must start with a non-digit character or underscore.

The line with the closing identifier can contain no other characters, except possibly a semicolon (;). The identifier may not be indented, and there may not be any spaces or tabs after or before the semicolon. Further, the first character before the closing identifier must be a newline as defined by your operating system. (This is \r on Macintosh for example.)

If this rule is broken and the closing identifier is not "clean" then it's not considered to be a closing identifier and PHP will continue looking for one. If in this case a proper closing identifier is not found then a parse error will result with the line number reported as being at the end of the script.

Now that we've introduced (most of) PHP's data types, let's go back to the boolean type and talk about PHP's notion of truth and falsehood.

Truth and falsehood according to PHP

For PHP, only the following are false:

- the boolean value FALSE itself;
- the empty string " ";
- the string "0";
- the integer 0;
- the float 0.0;
- the NULL value;
- an empty array (an array with no elements in it -- we will get to arrays later);
- an object with no properties set (hereagain, objects will come up later).

Absolutely everything else is TRUE, including negative numbers, the string "false", or anything else you can think of that you might intuitively assume to be false.

Operators

In this discussion we will run through the operators that you will use most frequently. For an exhaustive (and authoritative) treatment please see <http://us3.php.net/manual/en/language.operators.php>, from which much of the following is stolen.

The **precedence** of an operator specifies how "tightly" it binds two expressions together. For example, in the expression $1 + 5 * 3$, the answer is 16 and not 18 because the multiplication ("*") operator has a higher precedence than the addition ("+") operator. Parentheses may be used to force precedence. For instance: $(1 + 5) * 3$ evaluates to 18. If operator precedence is equal, left to right associativity is used to break the tie, so to speak.

associativity means the order in which PHP reads the expression -- either left to right, or right to left -- if precedence does not determine what gets evaluated first. For example, you might wonder whether the output of this would be 8 or -6:

```
<?php
    echo 4 - 3 + 7;
?>
```

Nobody who reads your code -- not even you -- wants to spend time pondering questions like the above, so a good rule of thumb is simply: always use parentheses, whether it's strictly necessary or not.

The **arithmetic operators** behave as you would expect: + - * / and % for addition, subtraction, multiplication, division, and modulus, respectively.

```
<?php $foo = 23; ?>
```

The above example assigns 23, the value on the right side (of the = operator), to the variable \$foo on the left. You can't do it the other way around because 23 is a literal value rather than a variable.

Note that an assignment expression itself evaluates to the value assigned. And since assignment associates from right to left, you will sometimes see things like

```
<?php  
    $a = $b = 23;  
?>
```

In addition to the basic assignment operator, there are "**combined operators**" for all of the binary arithmetic and string operators that allow you to use a value in an expression and then set its value to the result of that expression : += -= /= *= %=

You'll probably find the following example more comprehensible than the preceding sentence:

```
<?php  
    $foo = 3;  
    $foo += 2; // sets foo to 3 + 2, or 5  
    $foo *= 2; // sets foo to 5 * 2  
?>
```

Incrementing/Decrementing Operators: PHP (like other languages) provides the ++ and -- auto-increment and auto-decrement operators. They increase (or decrease) by the value of the variable that they are applied to. There are two flavors of each: pre- and post-. In the case of the former, the value is changed and assigned to the variable *before* it is evaluated in the expression where it appears; with the latter, the value changes afterwards. In both cases the value is increased/decreased by one; the difference is simply a matter of when. Example:

```
<?php  
    $x = $y = 1;  
    // increments y before the string concatenation  
    echo "value of y is " . ++$y . "\n";  
    // increments x after the string concatenation  
    echo "value of x is " . $x++ . "\n";  
?>
```

Output:

```
value of y is 2  
value of x is 1
```

The **comparison operators** are

```
$a == $b      // true if $a and $b are equal  
$a === $b    // true if $a and $b are equal AND the same data type  
$a != $b     // true if $a and $b are not equal  
$a <> $b     // same as above  
$a !== $b    // true if $a and $b are not equivalent OR not of the same type  
$a < $b      // true if $a is less than $b  
$a > $b      // true if $a is greater than $b  
$a <= $b     // true if $a is equal to or less than $b  
$a >= $b     // true if $a is equal to or greater than $b
```

If you compare an integer with a string, PHP automatically converts the string to a number. If you compare two numerical strings, they are compared as integers.

TIP: Beware of saying `=` when you mean `==`. Remember that an assignment expression itself evaluates to the value assigned. Remember also what we just noted about truth and falsehood. That means if you say 'assign' like so

```
<?php
    if ($a = 4) { // do something }
?>
```

when you really meant to say `==` ('compare'), your if-condition will always evaluate to true because 4 is considered true, regardless of whether `$a` was true or false before that point in your program. (If you don't understand this now, don't worry: you will eventually learn it the hard way (-:)

The **logical operators** are analogous to our everyday understanding of the words 'and' and 'or' and 'not':

or	The expression (<code>\$a or \$b</code>) is true if either <code>\$a</code> or <code>\$b</code> is true.
and	The expression (<code>\$a and \$b</code>) is true if both <code>\$a</code> and <code>\$b</code> are true.
! (not/negation)	The expression (<code>! \$a</code>) is true if <code>\$a</code> is false.
xor (exclusive or)	The expression (<code>\$a xor \$b</code>) is true if <i>either</i> <code>\$a</code> or <code>\$b</code> is true but <i>not both</i> .

There are also the `&&` and `||` operators that also mean 'and' and 'or', respectively, but have higher precedence than their English-like counterparts. (We will see an example later in which that fact allows you to write compact, clear and idiomatic code.) Here's an example of an expression involving 'or' and comparison operators:

```
<?php
    if ($temperature > 75 or $temperature < 60) {
        echo "I'm uncomfortable";
    } else {
        echo "I'm fine thanks";
    }
?>
```

There are just two **string operators** (the concatenation operator) and `.=` (concatenating assignment operator).

```
<?php
$string = "Hello";
$string = "Hello" . " "; // appends a space character to $string
$string .= "World";    // appends "World" to $string
echo $string;          // outputs "Hello World"
?>
```

Finally, there is the error suppression operator: `@`. It suppresses any error output from the expression that it precedes. We will understand this better later on when we talk about error handling.

Operators we are not talking about: because we will only very rarely use them in this class, for the sake of saving time we are skipping the **bitwCSE operators**. See <http://www.php.net/manual/en/language.operators.bitwCSE.php> if you are truly curious.

Control structures

If you are already familiar with other program languages, you will probably find PHP's control structures familiar. If on the other hand PHP is your first language, then you'll find this knowledge helpful if you ever decide to study another language.

Control flow structures manage the flow of execution of your program. They (along with *variables*)

The line with the closing identifier can contain no other characters, except possibly a semicolon (;). The identifier may not be indented, and there may not be any spaces or tabs after or before the semicolon. Further, the first character before the closing identifier must be a newline as defined by your operating system. (This is \r on Macintosh for example.)

If this rule is broken and the closing identifier is not "clean" then it's not considered to be a closing identifier and PHP will continue looking for one. If in this case a proper closing identifier is not found then a parse error will result with the line number reported as being at the end of the script.

Now that we've introduced (most of) PHP's data types, let's go back to the boolean type and talk about PHP's notion of truth and falsehood.

Truth and falsehood according to PHP

For PHP, only the following are false:

- the boolean value FALSE itself;
- the empty string " ";
- the string "0";
- the integer 0;
- the float 0.0;
- the NULL value;
- an empty array (an array with no elements in it -- we will get to arrays later);
- an object with no properties set (hereagain, objects will come up later).

Absolutely everything else is TRUE, including negative numbers, the string "false", or anything else you can think of that you might intuitively assume to be false.

if

`if` is perhaps the most fundamental construct in computer programming: it allows for conditional execution of a fragment of code. In pseudocode, the syntax is:

```
if (expression) {
    statement
}
```

If `expression` is true, `statement` is executed; otherwise it is ignored, and the flow of execution drops through to whatever comes after the `if` block. You can have multiple statements, and you can nest `if` blocks as deep as desired.

Note the curly braces. They are optional for single line statements (but in my opinion omitting them is a bad habit). However, they must be paired, i.e., you need a left curly and a right curly, or you will have problems.

Also, note the indentation of `statement`. Although PHP does not require it, indenting your blocks aids legibility immensely; not indenting is considered poor style. These remarks apply to all the control structures, not just `if`.

This is important enough to bear repeating: indent your blocks. It will save you a world of pain.

else

`else` extends an `if` statement to execute a statement in case the expression in the `if` statement evaluates to FALSE. For example, the following code would display "a is bigger than b" if `$a` is bigger than `$b`, and "a is NOT bigger than b" otherwise.

```
<?php
if ($a > $b) {
    echo "a is bigger than b";
} else {
    echo "a is NOT bigger than b";
}
?>
```

elseif

`elseif`, as its name suggests, is a combination of `if` and `else`. Like `else`, it extends an `if` statement to execute a different statement in case the original `if` expression evaluates to FALSE. However, unlike `else`, it will execute that alternative expression only if the `elseif` conditional expression evaluates to TRUE. For example, the following code would display a is bigger than b, a equal to b or a is smaller than b:

```
<?php
if ($a > $b) {
    echo "a is bigger than b";
} elseif ($a == $b) {
    echo "a is equal to b";
} else {
    echo "a is smaller than b";
}
?>
```

while

`while` is for *looping*, that is, repeating some operation as long as some condition is true. In pseudocode, the syntax is:

```
while (expression)
    statement
```

`while` can be thought of as similar to `if`, except that instead of executing once, it keeps executing `statement` unless and until `expression` is false. If `expression` is false the first time it is evaluated, the code is executed zero times.

What do you suppose the output of the following fragment would be?

```
<?php
$i = 1;
while ($i <= 10) {
    echo $i++;
}
?>
```

If you said "12345678910" you'd be right.

do-while

`do-while` loops are similar to `while` loops, except the truth expression is checked at the end of each iteration instead of at the beginning. The main difference from regular `while` loops is that the first iteration of a `do-while` loop is guaranteed to run (the truth expression is only checked at the end of the iteration), whereas it may not necessarily run with a regular `while` loop (the truth expression is checked at the beginning of each iteration; if it evaluates to FALSE right from the beginning, the loop execution would end immediately).

```
<?php
$i = 0;
do {
    echo $i;
} while ($i > 0);
?>
```

You can get usually by without `do-while`; it is not frequently used.

for

`for` is the most complex of the PHP looping constructs, and it is frequently used.

```
for (expr1; expr2; expr3)
    statement
```

The first expression (`expr1`) is evaluated (executed) once unconditionally at the beginning of the loop. In the beginning of each iteration, `expr2` is evaluated. If it evaluates to TRUE, the loop continues and the nested `statement(s)` are executed. If it evaluates to FALSE, the execution of the loop ends. At the end of each iteration, `expr3` is evaluated (executed).

`for` is often used for keeping track of counters, like so:

```
<?php
```

```
for ($i = 1; $i <= 10; $i++) {
    echo $i;

while (there are more job applicants) {

    // bring in the next applicant for an interview

    if (this applicant is a total loser) {
        continue; // get out of my office. next, please?
    }

    if (this applicant is a genius)  {
        break; // stop interviewing people, this person is hired
    }

    // if neither of above conditions is met, execution gets to this point.
    // carry on interview with current applicant.
}
```

\$i is initialized to 1; \$i is compared to 10; \$i is incremented if and only if \$i is equal to or less than 10.

There is more to be said about `for`, but the simple usage illustrated above is the most common. Please see <http://us3.php.net/manual/en/control-structures.for.php> if you're curious about the fine print.

foreach

`foreach` is for iterating through arrays. We will discuss `foreach` when we discuss arrays.

break and **continue**

`break` immediately ends execution of the current `for`, `foreach`, `while`, `do-while` or `switch` structure (we haven't yet discussed `switch`) . You use `break` whenever you are inside a loop and want to get out now for some reason, and move on to execution of whatever follows the loop structure.

`continue` is used within looping structures to skip the rest of the current loop iteration and continue execution at the beginning of the next iteration. In other words, you `continue` when you want to abort the current iteration and go on to the next iteration immediately.

A pseudo-code example might elucidate. Suppose you are a manager interviewing a series of job applicants for a position:

```
while (there are more job applicants) {  
    // bring in the next applicant for an interview  
  
    if (this applicant is a total loser) {  
        continue; // get out of my office. next, please?  
    }  
  
    if (this applicant is a genius) {  
        break; // stop interviewing people, this person is hired  
    }  
  
    // if neither of above conditions is met, execution gets to this point.  
    // carry on interview with current applicant.  
}
```

Unit – 8

Rails and Ruby

8.1 Origins and Uses of Ruby

Ruby was designed in Japan by Yukihiro Matsumoto (a.k.a. Matz) and was released in 1996. It started as a replacement for Perl and Python, languages that Matz found inadequate for his purposes. The use of Ruby grew rapidly in Japan and spread to the rest of the world a few years later. The quick growth of the use of Rails, the Web application development framework that is written in Ruby and uses Ruby, has accelerated the expansion of the language. Rails is probably the most common use of Ruby.

Learning Ruby is made easier by its implementation method: pure interpretation. Rather than needing to learn about and write a layer of boilerplate code around some simple logic, in Ruby one can write just that simple logic and request its interpretation. For example, consider the difference between a complete “Hello, World” program in a language like C++ or Java and the Ruby “Hello, World” program:

```
puts "Hello, World"
```

From Perl, Ruby gets regular expressions and implicit variables. From Java-Script, it gets objects that can change during execution. However, Ruby has many more differences with those languages than it has similarities. For example, as in pure object-oriented languages, every data value in Ruby is an object, whether it is a simple integer literal or a complete file system.

Ruby is available for every common computing platform. Furthermore, as is the case with PHP, the Ruby implementation is free.

8.2 Scalar Types and Their Operations

Ruby has three categories of data types: scalars, arrays, and hashes. The most commonly used type: scalars. There are two categories of scalar types: numeric's and character strings.

Numeric and String Literals

All numeric data types in Ruby are descendants of the Numeric class. The immediate child classes of Numeric are Float and Integer. The Integer class has two child classes: Fixnum and Bignum.

An integer literal that fits into the range of a machine word, which is often 32 bits, is a Fixnum object. An integer literal that is outside the Fixnum range is a Bignum object. Though it is odd among programming languages, there is no length limitation (other than your computer's memory size) on integer literals. If a Fixnum integer grows beyond the size limitation of Fixnum objects, it is coerced to a Bignum object. Likewise, if an operation on a Bignum object results in a value that fits into a Fixnum object, it is coerced to a Fixnum type.

Underscore characters can appear embedded in integer literals. Ruby ignores such underscores, allowing large numbers to be slightly more readable. For example, instead of 124761325, 124_761_325 can be used.

A numeric literal that has either an embedded decimal point or a following exponent is a Float object, which is stored as the underlying machine's double-precision floating-point type. The decimal point must be embedded; that is, it must be both preceded and followed by at least one digit. So, .435 is not a legal literal in Ruby.

All string literals are String objects, which are sequences of bytes that represent characters. There are two categories of string literals: single quoted and double quoted. Single-quoted string literals cannot include characters specified with escape sequences, such as newline characters specified with \n. If an actual single-quote character is needed in a string literal that is delimited by single quotes, the embedded single quote is preceded by a backslash, as in the following example:

'I'll meet you at O\'Malleys'

If an escape sequence is embedded in a single-quoted string literal, each character in the sequence is taken literally as itself. For example, the sequence \n in the following string literal will be treated as two characters—a backslash and an n:

'Some apples are red, \n some are green'

If a string literal with the same characteristics as a single-quoted string is needed, but you want to use a different delimiter, precede the delimiter with q, as in the following example:

q\$Don't you think she's pretty?\$

If the new delimiter is a parenthesis, a brace, a bracket, or an angle bracket, the left element of the pair must be used on the left and the right element must be used on the right, as in

q<Don't you think she's pretty?>

Double-quoted string literals differ from single-quoted string literals in two ways: First, they can include special characters specified with escape sequences; second, the values of variable names can be interpolated into the string, which means that their values are substituted for their names.

In many situations, special characters that are specified with escape sequences must be included in string literals. For example, if the words on a line must be spaced by tabs, a double-quoted literal with embedded escape sequences for the tab character can be used as in the following string:

"Runs \t Hits \t Errors"

A double quote can be embedded in a double-quoted string literal by preceding it with a backslash.

A different delimiter can be specified for string literals with the characteristics of double-quoted strings by preceding the new delimiter with Q as follows:

Q@”Why not learn Ruby?”, he asked. @

The null string (the string with no characters) can be denoted with either ““ or ““.

Variables and Assignment Statements

Naming conventions in Ruby help identify different categories of variables. For now, we will deal with local variables only. Other naming conventions will be explained as needed.

The form of variable names is a lowercase letter or an underscore, followed by any number of uppercase or lowercase letters, digits, or underscores. The letters in a variable name are case sensitive, meaning that fRIZZY, frizzy, frIzZy, and frizzY are all distinct names. However, by convention, programmer-defined variable names do not include uppercase letters.

As mentioned earlier, double-quoted string literals can include the values of variables. In fact, the results of executing any Ruby code can be included. This is specified by placing the code in braces and preceding the left brace with a pound sign (#). For example, if the value of tue_high is 83, then the string

“Tuesday’s high temperature was #{tue_high}”

has the following value:

“Tuesday’s high temperature was 83”

Similarly, if the value of price is 1.65 and that of quantity is 6, then the value of the string

“The cost of our apple order is \$#{price * quantity}”

Is “The cost of our apple order is \$9.90”

Because Ruby is a pure object-oriented programming language, all of its variables are references to objects. This is in contrast to more conventional languages, such as C++ and Java, which have two categories of variables: those for primitives and those that reference objects. In Ruby, every data value is an object, so it needs references only. Because references are typeless, there is no point in declaring them. In fact, there is no way to declare a variable in Ruby.

A scalar variable that has not been assigned a value by the program has the value nil.

Ruby has constants, which are distinguished from variables by their names, which always begin with uppercase letters. A constant is created when it is assigned a value, which can be any constant expression. In Ruby, a constant can be assigned a new value, although it causes a warning message to the user.

Ruby includes some predefined, or implicit, variables. The name of an implicit scalar variable begins with a dollar sign. The rest of the name is often just one more special character, such as an underscore (_), a circumflex (^), or a backslash (\). This chapter and the next include some uses of these implicit variables.

The assignment statements of Ruby are exactly like those of the programming languages derived from C. The only thing to remember is that the variables of Ruby are all typeless references. All that is ever assigned in an assignment statement is the address of an object.

Numeric Operators

Most of Ruby's numeric operators are similar to those in other common programming languages, so they should be familiar to most readers. There are the binary operators: + for addition, - for subtraction, * for multiplication, / for division, ** for exponentiation, and % for modulus. The modulus operator is defined as follows: $x \% y$ produces the remainder of the value of x after division by y . If an integer is divided by an integer, integer division is done. Therefore, $3 / 2$ produces 1.

The precedence rules of a language specify which operator is evaluated first when two operators that have different levels of precedence appear in an expression and are separated only by an operand. The associativity rules of a language specify which operator is evaluated first when two operators with the same precedence level appear in an expression and are separated only by an operand. The precedence and associativity of the numeric operators are given in table below:

Operator	Associativity
**	Right
unary +, -	Right
*, /, %	Left
binary +, -	Left

The operators listed first have the highest precedence.

Note that Ruby does not include the increment (++) and decrement (--) operators found in all of the C-based languages.

Ruby includes the Math module, which has methods for basic trigonometric and transcendental functions. Among these methods are cos (cosine), sin (sine), log (logarithm), sqrt (square root), and tan (tangent). The methods of the Math module are referenced by prefixing their names with Math., as in Math.sin(x). All of these take any numeric type as a parameter and return a Float value.

Included with the Ruby implementation is an interactive interpreter, which is very useful to the student of Ruby. It allows one to type any Ruby expression and get an immediate response from the interpreter. The interactive interpreter's name is Interactive Ruby, whose acronym, IRB, is the name of the program that supports it. One enters irb simply by typing irb at the command prompt in the directory that contains the Ruby interpreter. For example, if the command prompt is a percent sign (%), one can type

```
% irb
```

after which irb will respond with its own prompt, which is

```
irb(main):001:0>
```

At this prompt, any Ruby expression or statement can be typed, whereupon irb interprets the expression or statement and returns the value after an implication symbol (\Rightarrow), as in the following example:

```
irb(main):001:0> 17 * 3
```

```
=> 51
```

```
irb(main):002:0>
```

The lengthy default prompt can be easily changed. We prefer the simple “ $>>>$ ” prompt. The default prompt can be changed to this with the following command:

```
irb(main):002:0> conf.prompt_i = ">>>"
```

From here on, we will use this simple prompt.

String Methods

The Ruby String class has more than 75 methods, a few of which are described in this section. Many of these methods can be used as if they were operators. In fact, we sometimes call them operators, even though underneath they are all methods.

The String method for catenation is specified by plus (+), which can be used as a binary operator. This method creates a new string from its operands:

```
>> "Happy" + " " + "Holidays!"
```

```
=> "Happy Holidays!"
```

The $<<$ method appends a string to the right end of another string, which, of course, makes sense only if the left operand is a variable. Like +, the $<<$ method can be used as a binary operator. For example, in the interactions

```
>> mystr = "G'day,"  
=> "G'day,"  
>> mystr << "mate"  
=> "G'day, mate"
```

the first assignment creates the specified string literal and sets the variable mystr to reference that memory location. If mystr is assigned to another variable, that variable will reference the same memory location as mystr:

```
>> mystr = "Wow!"
=> "Wow!"
>> yourstr = mystr
=> "Wow!"
>> yourstr
=> "Wow!"
```

Now both mystr and yourstr reference the same memory location: the place that has the string “Wow!”. If a different string literal is assigned to mystr, Ruby will build a memory location with the value of the new string literal and mystr will reference that location.

The other most commonly used methods of Ruby are similar to those of other programming languages. Among these are the ones shown in Table below; all of them create new strings.

Method	Action
<code>capitalize</code>	Converts the first letter to uppercase and the rest of the letters to lowercase
<code>chop</code>	Removes the last character
<code>chomp</code>	Removes a newline from the right end if there is one
<code>upcase</code>	Converts all of the lowercase letters in the object to uppercase
<code>downcase</code>	Converts all of the uppercase letters in the object to lowercase
<code>strip</code>	Removes the spaces on both ends
<code>lstrip</code>	Removes the spaces on the left end
<code>rstrip</code>	Removes the spaces on the right end
<code>reverse</code>	Reverses the characters of the string
<code>swapcase</code>	Converts all uppercase letters to lowercase and all lowercase letters to uppercase

As stated previously, all of these methods produce new strings, rather than modify the given string in place. However, all of the methods also have versions that modify their objects in place. These methods are called bang or mutator methods and are specified by following their names with an exclamation point (!). To illustrate the difference between a string method and its bang counterpart, consider the following interactions:

```
>> str = "Frank"
=> "Frank"
>> str.upcase
=> "FRANK"
>> str
=> "Frank"
>> str.upcase!
=> "FRANK"
>> str
=> "FRANK"
```

Note that, after upcase is executed, the value of str is unchanged (it is still “Frank”), but after upcase! is executed, it is changed (it is “FRANK”).

Ruby strings can be indexed, somewhat as if they were arrays. As one would expect, the indices begin at zero. The brackets of this method specify a getter method. The catch is that the getter method returns the ASCII code (as a Fixnum object), rather than the character. To get the character, the chr method must be called, as in the following interactions:

```
>> str = "Shelley"
=> "Shelley"
>> str[1]
=> 104
>> str[1].chr
=> "h"
```

If a negative subscript is used as an index, the position is counted from the right.

A multicharacter substring of a string can be accessed by including two numbers in the brackets, in which case the first is the position of the first character of the substring and the second is the number of characters in the substring. Unlike the single-character reference, however, in this case the value is a string, not a number:

```
>> str = "Shelley"
=> "Shelley"
>> str[2,4]
=> "elle"
```

The substring getter can be used on individual characters to get one character without calling the chr method. Specific characters of a string can be set with the setter method, []=, as in the following interactions:

```
>> str = "Donald"
=> "Donald"
>> str[3,3] = "nie"
=> "nie"
>> str
=> "Donnie"
```

The usual way to compare strings for equality is to use the `==` method as an operator:

```
>> "snowstorm" == "snowstorm"
=> true
>> "snowie" == "snowy"
=> false
```

A different sense of equality is tested with the `equal?` method, which determines whether its parameter references the same object as the one to which it is sent. For example, the interactions

```
>> "snowstorm".equal?("snowstorm")
=> false
```

produces `false` because, although the contents of the two string literals are the same, they are different objects.

Yet another sense of equality is tested with the `eql?` method, which returns `true` if its receiver object and its parameter have the same types and the same values. The following interactions illustrate an instance of equality and an instance of inequality:

```
>> 7 == 7.0
=> true
>> 7.eql?(7.0)
=> false
```

To facilitate ordering, Ruby includes the “spaceship” operator, `<=>`, which returns `-1` if the second operand is greater than the first, `0` if the two operands are equal, and `1` if the first operand is greater than the second. “Greater” in this case means that the text in question belongs later alphabetically. The following interactions illustrate all three cases:

```
>> "apple" <=> "prune"
=> -1
>> "grape" <=> "grape"
=> 0
```

```
>> "grape" <=> "apple"
```

```
=> 1
```

The repetition operator is specified with an asterisk (*). It takes a string as its left operand and an expression that evaluates to a number as its right operand. The left operand is replicated the number of times equal to the value of the right operand:

```
>> "More! " * 3
```

```
=> "More! More! More! "
```

8.3 Simple Input and Output

Screen Output

Output is directed to the screen with the puts method (or operator). We prefer to treat it as an operator. The operand for puts is a string literal. A newline character is implicitly appended to the string operand. If the value of a variable is to be part of a line of output, the #{...} notation can be used to insert it into a double-quoted string literal, as in the following interactions:

```
>> name = "Fudgy"
```

```
=> "Fudgy"
```

```
>> puts "My name is #{name}"
```

```
My name is Fudgy
```

```
=> nil
```

The value returned by puts is nil, and that is the value returned after the string has been displayed.

The print method is used if you do not want the implied newline that puts adds to the end of your literal string.

The way to convert a floating-point value to a formatted string is with a variation of the C language function sprintf. This function, which also is named sprintf, takes a string parameter that contains a format code followed by the name of a variable to be converted. The string version is returned by the function. The format codes most commonly used are f and d. The form of a format code is a percent sign (%), followed by a field width, followed by the code letter (f or d). The field width for the f code appears in two parts, separated by a decimal point. For example, %f7.2 means a total field width of seven spaces, with two digits to the right of the decimal point—a perfect format for money. The d code field width is just a number of spaces—for example, %5d. So, to convert a floating-point value referenced by the variable total to a string with two digits to the right of the decimal point, the following statement could be used:

```
str = sprintf("%5.2f", total)
```

Because Ruby is used primarily for Rails in this book, there is little need for keyboard input. However, keyboard input is certainly useful for other applications, so it is briefly introduced here.

The gets method gets a line of input from the keyboard. The retrieved line includes the newline character. If the newline is not needed, it can be discarded with chomp:

```
>> name = gets
```

```
apples
```

```
=> "apples\n"
```

```
>> name = name.chomp
```

```
=> "apples"
```

This code could be shortened by applying chomp directly to the value returned by gets:

```
>> name = gets.chomp
```

```
apples
```

```
=> "apples"
```

If a number is to be input from the keyboard, the string from gets must be converted to an integer with the to_i method, as in the following interactions:

```
>> age = gets.to_i
```

```
27
```

```
=> 27
```

If the number is a floating-point value, the conversion method is to_f:

```
>> age = gets.to_f
```

```
27.5
```

```
=> 27.5
```

In this same vein, we must mention that there is a similar method, to_s, to which every object responds. The method converts the value of the object to which it is sent to a string. However, because puts implicitly converts its operand to a string, to_s is not often explicitly called.

The following listing is of a trivial program created with a text editor and stored in a file:

```
# quadeval.rb - A simple Ruby program
# Input: Four numbers, representing the values of
#         a, b, c, and x
# Output: The value of the expression
#         a*x**2 + b*x + c
# Get input
puts "Please input the value of a "
a = gets.to_i
puts "Please input the value of b "
b = gets.to_i
puts "Please input the value of c "
c = gets.to_i
puts "Please input the value of x "
x = gets.to_i
# Compute and display the result
result = a * x ** 2 + b * x + c
puts "The value of the expression is: #{result}"
```

A program stored in a file can be run by the command

>ruby filename

So, our example program can be run (interpreted) with

>ruby quadeval.rb

To compile, but not interpret, a program, just to check the syntactic correctness of the program, the -c flag is included after the ruby command. It is also a good idea to include the -w flag, which causes ruby to produce warning messages for a variety of suspicious things it may find in a program. For example, to check the syntax of our example program, the following statement could be used:

>ruby -cw quadeval.rb

If the program is found to be syntactically correct, the response to this command is as follows:

Syntax OK

8.4 Control Statements

Ruby includes a complete collection of statements for controlling the flow of execution through programs. This section introduces the control expressions and control statements of Ruby.

The expressions upon which statement control flow is based are Boolean expressions. They can be either of the constants true or false, variables, relational expressions, or compound expressions. A control expression that is a simple variable is true if its value is anything except nil (in other words, if it references some object). If its value is nil, it is false.

A relational expression has two operands and a relational operator. Relational operators can have any scalar-valued expression as operands. The relational operators of Ruby are shown in Table below:

Operator	Operation
<code>==</code>	Is equal to
<code>!=</code>	Is not equal to
<code><</code>	Is less than
<code>></code>	Is greater than
<code><=</code>	Is less than or equal to
<code>>=</code>	Is greater than or equal to
<code><=></code>	Compare, returning -1, 0, or +1
<code>eql?</code>	True if the receiver object and the parameter have the same type and equal values
<code>equal?</code>	True if the receiver object and the parameter have the same object ID

Recall that the `<=>` operator is often used for comparing strings. Also, `equal?` is used to determine whether two variables are aliases (i.e., whether they reference the same object).

Ruby has two sets of operators for the AND, OR, and NOT Boolean operations. The two sets have the same semantics but different precedence levels. The operators with the higher precedence are `&&` (AND), `||` (OR), and `!` (NOT). Those with the lower precedence are `and`, `or`, and `not`. The precedence of these latter operators is lower than that of any other operators in Ruby, so, regardless of what operators appear in their operands, these operators will be evaluated last.

All of the relational operators are methods, but all except `eql?` and `equal?` can be used as operators.

The precedence and associativity of all operators discussed so far in this chapter are shown in Table below:

Operator	Associativity
<code>**</code>	Right
<code>!, unary + and -</code>	Right
<code>*, /, %</code>	Left
<code>+, -</code>	Left
<code>&</code>	Left
<code>+, -</code>	Left
<code>>, <, >=, <=</code>	Nonassociative
<code>==, !=, <=></code>	Nonassociative
<code>&&</code>	Left
<code> </code>	Left
<code>=, +=, -=, *=, **=, /=, %=, &=, &&=, =</code>	Right
<code>not</code>	Right
<code>or, and</code>	Left

Because assignment statements have values (the value of an assignment is the value assigned to the left-side variable), they can be used as control expressions. One common such application is for an assignment statement that reads a line of input. The `gets` method returns `nil` when it gets the end-of-file (EOF) character, so this character can be conveniently used to terminate loops. A typical example is

```
while (next = gets) { ... }
```

The keyboard EOF character is Control-D in UNIX, Control-Z in Windows, and CMD+. (period) in Macintosh systems.

Selection and Loop Statements

Control statements require some syntactic container for sequences of statements whose execution they are meant to control. The Ruby form of such containers is to use a simple sequence of statements terminated with `else` (if the sequence is a `then` clause) or `end` (if the sequence is either an `else` clause or a `then` clause, in which case there is no `else` clause). A control construct is a control statement together with the segment of code whose execution it controls.

Ruby's `if` statement is similar to that of other languages. One syntactic difference is that there are no parentheses around the control expression, as is the case with most of the languages based directly or even loosely on C. The following construct is illustrative:

```
if a > 10
```

```
b = a * 2
```

```
end
```

An if construct can include elsif (note that it is not spelled “elseif”) clauses, which provide a way of having a more readable sequence of nested if constructs. The following if construct is typical:

```
if snowrate < 1
```

```
  puts "Light snow"
```

```
elsif snowrate < 2
```

```
  puts "Moderate snow"
```

```
else
```

```
  puts "Heavy snow"
```

```
end
```

Ruby has an unless statement, which is the same as its if statement, except that the inverse of the value of the control expression is used. This is convenient if you want a selection construct with an else clause but no then clause. The following construct illustrates an unless statement:

```
unless sum > 1000
```

```
  puts "We are not finished yet!"
```

```
end
```

Ruby includes two kinds of multiple selection constructs, both named case. One Ruby case construct, which is similar to a switch, has the following form:

```
case expression
when value then
  - statement sequence
...
when value then
  - statement sequence
[else
  - statement sequence]
end
```

The value of the case expression is compared with the values of the when clauses, one at a time, from top to bottom, until a match is found, at which time the sequence of statements that follow is interpreted. The comparison is done with the `==` relational operator, which is defined for all built-in classes. If the when value is a range, such as `(1..100)`, `==` is defined as an inclusive test, yielding true if the value of the case expression is in the given range. If the when value is a class

name, `==` is defined to yield true if the case value is an object of the case expression class or one of its superclasses. If the when value is a regular expression, `==` is defined to be a simple pattern match. Note that the `==` operator is used only for the comparisons in case constructs.

Consider the following example:

```
case in_val
when -1 then
  neg_count += 1
when 0 then
  zero_count += 1
when 1 then
  pos_count += 1
else
  puts "Error - in_val is out of range"
end
```

Note that no break statements are needed at the ends of the sequences of selectable statements in this construct: There are implied branches at the end of each when clause that cause execution to exit the construct.

```
case
when Boolean expression then expression
...
when Boolean expression then expression
else expression
end
```

The semantics of the construct is straightforward. The Boolean expressions are evaluated one at a time, until one evaluates to true. The value of the whole construct is the value of the expression that corresponds to the true Boolean expression. If none of the Boolean expressions is true, the else expression is evaluated and its value is the value of the construct. For example, consider the following assignment statement:

```
leap = case
  when year % 400 == 0 then true
  when year % 100 == 0 then false
  else year % 4 == 0
end
```

This case expression evaluates to true if the value of year is a leap year.

The Ruby while and for statements are similar to those of C and its descendants. The bodies of both are sequences of statements that end with `end`. The general form of the while statement is as

follows:

```
while control expression
      loop body statement(s)
end
```

The control expression could be followed by the do reserved word.

The until statement is similar to the while statement, except that the inverse of the value of the control expression is used.

For those situations in which a loop is needed in which the conditional termination is at some position in the loop other than the top, Ruby has an infinite loop construct and loop exit statements. The body of the infinite loop construct is like that of while: a sequence of statements that optionally begins with do and always ends with end.

There are two ways to control an infinite loop: with the break and next statements. These statements can be made conditional by putting them in the then clause of an if construct. The break statement causes control to go to the first statement following the loop body. The next statement causes control to go to the first statement in the loop body. For example, consider the following two infinite loop constructs:

```
sum = 0
loop do
    dat = gets.to_i
    if dat < 0 break
    sum += dat
end

sum = 0
loop do
    dat = gets.to_i
    if dat < 0 next
    sum += dat
end
```

In the first construct, the loop is terminated when a negative value is input. In the second, negative values are not added to sum, but the loop continues.

Ruby does not have a general for statement, which is ubiquitous among languages with C in their ancestry. However, Ruby includes convenient ways to construct the counting loops implemented with for statements in other common languages. These loops are built with iterator methods, which we postpone discussing until methods and arrays have been introduced. Also, there are the for and for-in constructs in Ruby, which are used for iterating through arrays and hashes (associative arrays).

8.5 Fundamentals of Arrays

Ruby includes two structured classes or types: arrays and hashes.

Arrays in Ruby are more flexible than those of most of the other common languages. This flexibility is a result of two fundamental differences between Ruby arrays and those of other common languages such as C, C++, and Java. First, the length of a Ruby array is dynamic: It can grow or shrink anytime during program execution. Second, a Ruby array can store different types of data. For example, an array may have some numeric elements, some string elements, and even some array elements. So, in these cases, Ruby arrays are similar to those of PHP.

Ruby arrays can be created in two different ways. First, an array can be created by sending the new message to the predefined Array class, including a parameter for the size of the array. The second way is simply to assign a list literal to a variable, where a list, literal is a list of literals delimited by brackets. For example, in the following interactions, the first array is created with new and the second is created by assignment:

```
>> list1 = Array.new(5)
=> [nil, nil, nil, nil, nil]
>> list2 = [2, 4, 3.14159, "Fred", [] ]
=> [2, 4, 3.14159, "Fred", []]
```

An array created with the new method can also be initialized by including a second parameter, but every element is given the same value (that of the second parameter). Thus, we may have the following interactions:

```
>> list1 = Array.new(5, "Ho")
=> ["Ho", "Ho", "Ho", "Ho", "Ho"]
```

Actually, this form of initialization is rarely useful, because not only is each element given the same value, but also each is given the same reference. Thus, all of the elements reference the same object. So, if one is changed, all are changed.

All Ruby array elements use integers as subscripts, and the lower bound subscript of every array is zero. Array elements are referenced through subscripts delimited by brackets ([]), which actually constitutes a getter method that is allowed to be used as a unary operator. Likewise, []= is a setter method. A subscript can be any numeric-valued expression. If an expression with a floating-point value is used as a subscript, the fractional part is truncated. The following interactions illustrate the use of subscripts to reference array elements:

```
>> list = [2, 4, 6, 8]
=> [2, 4, 6, 8]
>> second = list[1]
=> 4
>> list[3] = 9
=> 9
>> list
=> [2, 4, 6, 9]
>> list[2.999999]
=> 6
```

The length of an array can be retrieved with the length method, as illustrated in the following interactions:

```
>> len = list.length
=> 4
```

The for-in Statement

The for-in statement is used to process the elements of an array. For example, the following code computes the sum of all of the values in list:

```
>> sum = 0
=> 0
>> list = [2, 4, 6, 8]
=> [2, 4, 6, 8]
>> for value in list
>>     sum += value
>> end
=> [2, 4, 6, 8]
>> sum
=> 20
```

Notice that the interpreter's response to the for-in construct is to display the list of values assumed by the scalar variable.

The scalar variable in a for-in takes on the values of the list array, one at a time. Notice that the scalar does not get references to array elements; it gets the values. Therefore, operations on the scalar variable have no affect on the array, as illustrated in the following interactions:

```
>> list = [1, 3, 5, 7]
=> [1, 3, 5, 7]
>> for value in list
>>   value += 2
>> end
=> [1, 3, 5, 7]
>> list
=> [1, 3, 5, 7]
```

A literal array value can be used in the for-in construct:

```
>> list = [2, 4, 6]
=> [2, 4, 6]
>> for index in [0, 1, 2]
>> puts "For index = #{index}, the value is #{list[index]}"
>> end
For index = 0, the element is 2
For index = 1, the element is 4
For index = 2, the element is 6
```

Built-In Methods for Arrays and Lists

This section introduces a few of the many built-in methods that are part of Ruby.

Frequently, it is necessary to place new elements on one end or the other of an array. Ruby has four methods for this purpose: unshift and shift, which deal with the left end of arrays; and pop and push, which deal with the right end of arrays.

The shift method removes and returns the first element (the one with lowest subscript) of the array object to which it is sent. For example, the following statement removes the first element of list and places it in first:

```
>> list = [3, 7, 13, 17]
=> [3, 7, 13, 17]
>> first = list.shift
=> 3
>> list
=> [7, 13, 17]
```

The subscripts of all of the other elements in the array are reduced by 1 as a result of the shift operation.

The pop method removes and returns the last element from the array object to which it is sent. In this case, there is no change in the subscripts of the array's other elements.

The unshift method takes a scalar or an array literal as a parameter and appends it to the beginning of the array. This requires an increase in the subscripts of all other array elements to create space in the array for the new elements. The push method takes a scalar or an array literal and adds it to the high end of the array:

```
>> list = [2, 4, 6]
```

```
=> [2, 4, 6]
```

```
>> list.push(8, 10)
```

```
=> {2, 4, 6, 8, 10}
```

Either pop and unshift or push and shift can be used to implement a queue in an array, depending on the direction in which the queue should grow.

Although push is a convenient way to add literal elements to an array, if an array is to be catenated to the end of another array, another method, concat, is used:

```
>> list1 = [1, 3, 5, 7]
=> [1, 3, 5, 7]
>> list2 = [2, 4, 6, 8]
=> [2, 4, 6, 8]
>> list1.concat(list2)
=> [1, 3, 5, 7, 2, 4, 6, 8]
```

If two arrays need to be catenated and the result saved as a new array, the plus (+) method can be used as a binary operator, as in the following interactions:

```
>> list1 = [0.1, 2.4, 5.6, 7.9]
=> [0.1, 2.4, 5.6, 7.9]
>> list2 = [3.4, 2.1, 7.5]
=> [3.4, 2.1, 7.5]
>> list3 = list1 + list2
=> [0.1, 2.4, 5.6, 7.9, 3.4, 2.1, 7.5]
```

Note that neither list1 nor list2 is affected by the plus method.

The reverse method does what its name implies:

```
>> list = [2, 4, 8, 16]
=> [2, 4, 8, 16]
>> list.reverse
=> [16, 8, 4, 2]
>> list
=> [2, 4, 8, 16]
```

Note that reverse returns a new array and does not affect the array to which it is sent. The mutator version of reverse, reverse!, does what reverse does, but changes the object to which it is sent:

```
>> list = [2, 4, 8, 16]
=> [2, 4, 8, 16]
>> list.reverse!
=> [16, 8, 4, 2]
>> list
=> [16, 8, 4, 2]
```

The include? predicate method searches an array for a specific object:

```
>> list = [2, 4, 8, 16]
=> [2, 4, 8, 16]
>> list.include?(4)
=> true
>> list.include?(10)
=> false
```

The sort method sorts the elements of an array, as long as Ruby is able to compare those elements. The most commonly sorted elements are either numbers or strings, and Ruby can compare numbers with numbers and strings with strings. So, sort works well on arrays of elements of either of these two types:

```
>> list = [16, 8, 4, 2]
=> [16, 8, 4, 2]
>> list.sort
=> [2, 4, 8, 16]
>> list2 = ["jo", "fred", "mike", "larry"]
=> ["jo", "fred", "mike", "larry"]
>> list2.sort
=> ["fred", "jo", "larry", "mike"]
```

If the sort method is sent to an array that has mixed types, Ruby produces an error message indicating that the comparison failed:

```
>> list = [2, "jo", 8, "fred"]
=> [2, "jo", 8, "fred"]
>> list.sort
ArgumentError: comparison of Fixnum with String failed
    from (irb):13:in 'sort'
    from (irb):13
    from :0
```

Note that sort returns a new array and does not change the array to which it is sent. By contrast, the mutator method, sort!, sorts the array to which it is sent, in place.

In some situations, arrays represent sets. There are three methods that perform set operations on two arrays: &, for set intersection; -, for set difference; and |, for set union. All are used as binary infix operators, as in the following examples:

```
>> set1 = [2, 4, 6, 8]
=> [2, 4, 6, 8]
>> set2 = [4, 6, 8, 10]
=> [4, 6, 8, 10]
>> set1 & set2
=> [4, 6, 8]
>> set1 - set2
=> [2]
>> set1 | set2
=> [2, 4, 6, 8, 10]
```

An Example

The example that follows illustrates a simple use of an array. A list of names is read from the keyboard. Each name is converted to all uppercase letters and placed in an array. The array is then sorted and displayed. Here is the document:

```
# process_names.rb - A simple Ruby program to
#   illustrate the use of arrays
#   Input: A list of lines of text, where each line
#         is a person's name
# Output: The input names, after all letters are
#         converted to uppercase, in alphabetical order

index = 0
names = Array.new

# Loop to read the names and process them
while (name = gets)

  # Convert the name's letters to uppercase and put it
  # in the names array
  names[index] = name.chomp.upcase
  index += 1
end

# Sort the array in place and display it
names.sort!
puts "The sorted array"
for name in names
  puts name
end
```

8.6 Hashes

Associative arrays are arrays in which each data element is paired with a key, which is used to identify the data element. Because hash functions are used both to create and to find specific elements in an associative array, associative arrays often are called hashes. There are two fundamental differences between arrays and hashes: First, arrays use numeric subscripts to address specific elements, whereas hashes use string values (the keys) to address elements; second, the elements in arrays are ordered by subscript, but the elements in hashes are not. In a sense, elements of an array are like those in a list, whereas elements of a hash are like those in a set, where order is irrelevant. The actual arrangement of the elements of a hash in memory is determined by the hash function used to insert and access them.

Like arrays, hashes can be created in two ways, with the new method or by assigning a literal to a variable. In the latter case, the literal is a hash literal, in which each element is specified by a key-value pair, separated by the symbol =>. Hash literals are delimited by braces, as in the following interactions:

```
>> kids_ages = {"John" => 41, "Genny" => 39, "Jake" => 25,
  "Darcie" => 24}
=> {"Darcie"=>24, "John"=>41, "Genny"=>39, "Jake"=>25}
```

Notice that the order of the hash returned by Ruby is not the same as the order in the hash literal

used to create the hash. This is because the actual order of the hash in memory is unpredictable (at least for the user program).

If the new method is sent to the Hash class without a parameter, it creates an empty hash, denoted by {}:

```
>> my_hash = Hash.new
```

```
=> {}
```

An individual value element of a hash can be referenced by “subscripting” the hash name with a key. The same brackets used for array element access are used to specify the subscripting operation:

```
>> kids_ages["Genny"]
```

```
=> 39
```

A new value is added to a hash by assigning the value of the new element to a reference to the key of the new element, as in the following example:

```
>> kids_ages["Aidan"] = 10;  
=> {"Aidan"=>10, "Darcie"=>24, "John"=>41, "Genny"=>39,  
"Jake"=>25}
```

An element is removed from a hash with the delete method, which takes an element key as a parameter:

```
>> kids_ages.delete("Genny")  
=> 39  
>> kids_ages  
=> {"Aidan"=>10, "Darcie"=>24, "John"=>41, "Jake"=>25}
```

A hash can be set to empty in one of two ways: either an empty hash literal can be assigned to the hash, or the clear method can be used on the hash. These two approaches are illustrated with the following statements:

```
>> hi_temps = {"mon" => 74, "tue" => 78}  
=> {"mon"=>74, "tue"=>78}  
>> hi_temps = {}  
=> {}  
>> salaries = {"Fred" => 47400, "Mike" => 45250}  
=> {"Fred" => 47400, "Mike" => 45250}  
>> salaries.clear  
=> {}
```

The has_key? predicate method is used to determine whether an element with a specific key is in a hash. The following interactions are illustrative, assuming that the kids_ages hash previously defined is still around:

```
>> kids_ages.has_key?("John")
=> true
>> kids_ages.has_key?("Henry")
=> false
```

The keys and values of a hash can be extracted into arrays with the methods keys and values, respectively:

```
>> kids_ages.keys
=> ["Aidan", "Darcie", "John", "Jake"]
>> kids_ages.values
=> [10, 24, 41, 25]
```

8.7 Methods

Subprograms are central to the usefulness of any programming language. Ruby's subprograms are all methods because it is an object-oriented language. However, Ruby's methods can be defined outside user-defined classes, so they are like functions, both in appearance and in behavior, when defined outside a class. When a method that is defined in a class is called from outside that class, the call must begin with a reference to an object of that class. When a method is called without an object reference, the default object on which it is called is self, which is a reference to the current object. Therefore, whenever a method is defined outside a user-defined class, it is called without an object reference.

Fundamentals

A method definition includes the method's header and a sequence of statements, ending with the end reserved word, that describes its actions. A method header is the reserved word def, the method's name, and optionally a parenthesized list of formal parameters. Method names must begin with lowercase letters. If the method has no parameters, the parentheses are omitted. In fact, the parentheses are optional in all cases, but it is common practice to include them when there are parameters and omit them when there are no parameters. The types of the parameters are not specified in the parameter list, because Ruby variables do not have types—they are all references to objects. The type of the return object is also not specified in a method definition.

A method that returns an object that is to be used immediately is called in the position of an operand in an expression (or as the whole expression). A method that does not return an object that is to be used can be called by a stand-alone statement.

A method can specify the value it returns in two ways: explicitly and implicitly. The return statement takes an expression as its parameter. The value of the expression is returned when the

return is executed. A method can have any number of return statements, including none. If there are no return statements in a method or if execution arrives at the end of the method without encountering a return, the object that is implicitly returned is the value of the last expression evaluated in the method.

The Time object is used to obtain various aspects of time from the system clock. The now method of Time returns the current time and date as a string. This method is used in the following example methods, one with a return and one without a return:

```
def date_time1
    return Time.now
end
def date_time2
    Time.now
end
```

The following calls to date_time1 and date_time2 yield the returned values shown:

```
>> date_time1
=> Thu Jun 07 16:00:06 Mountain Daylight Time 2007
>> date_time2
=> Thu Jun 07 16:00:08 Mountain Daylight Time 2007
```

Local Variables

Local variables either are formal parameters or are variables created in a method. A variable is created in a method by assigning an object to it. The scope of a local variable is from the header of the method to the end of the method. If the name of a local variable conflicts with that of a global variable, the local variable is used. This is the advantage of local variables: When you make up their names, you do not need to be concerned that a global variable with the same name may exist in the program.

The name of a local variable must begin with either a lowercase letter or an underscore (_). Beyond the first character, local variable names can have any number of letters, digits, or underscores.

The lifetime of a variable is the period over which it exists and can be referenced. The lifetime of a local variable is from the time it is created until the end of the execution of the method. So, the local variables of a method cannot be used to store data between calls to the method.

Parameters

The parameter values that appear in a call to a method are called actual parameters. The parameter names used in the method, which correspond to the actual parameters, are called formal parameters. In effect, scalar actual parameters specify the values of objects, not their addresses. So

in Ruby, the transmission of scalar parameters is strictly one way into the method. The values of the scalar actual parameters are available to the method through its formal parameters. The formal parameters that correspond to scalar actual parameters are local variables that are initialized to reference new objects that have the values of the corresponding actual parameters. Whatever a method does to its formal parameters, it has no effect on the actual parameters in the calling program unit. The following example illustrates a method that does not change its parameters:

```
def side3(side1, side2)
    return Math.sqrt(side1 ** 2 + side2 ** 2)
end
```

Now, we illustrate a method that attempts to change its parameters. The intent of the following method was to interchange its parameters:

```
>> def swap(x, y)
>> t = x
>> x = y
>> y = t
>> end
=> nil
>> a = 1
>> b = 2
>> swap(a, b)
=> 1
>> a
=> 1
>> b
=> 2
```

So, you see that, although swap changes its formal parameters, the actual parameters a and b sent to it are unchanged.

Actual parameters that are arrays or hashes are, in effect, passed by reference, so it is a two-way communication between the calling program unit and the called method. For example, if an array is passed to a method and the method changes the array, the changes are reflected in the corresponding actual parameter in the caller.

Normally, a call to a method must have the same number of actual parameters as the number of formal parameters in the method's definition. A mismatch of these two numbers results in a runtime error. However, a method can be defined to take a variable number of parameters by defining it with a parameter that is preceded by an asterisk (*). Such a parameter is called an asterisk parameter. For example, the method

```
def fun1(*params)
```

...

End

can take any number of parameters, including none. The actual parameters that are passed are placed in the array named params (in this example). The asterisk parameter can be preceded by other parameters, in which case only those actual parameters that do not correspond to named formal parameters are placed in the array of parameters. For example, suppose fun2 is defined as follows:

```
def fun2(sum, list, length, *params)
```

...

End

Now, suppose fun2 is called with the following statement:

```
fun2(new_sum, my_list, len, speed, time, alpha)
```

Then the actual parameters speed, time, and alpha will be passed into the array params. Of course, the asterisk parameter must always appear at the end of the list of formal parameters. Any normal parameters that follow an asterisk parameter will always be ignored, because the asterisk parameter receives all remaining actual parameters.

Formal parameters can have default values, making their corresponding actual parameters optional. For example, consider the following skeletal method definition:

```
def lister(list, len = 100)
```

...

End

If this method is called with the following statement, the formal parameter len gets the value 50:

```
lister(my_list, 50)
```

But if it is called with the following statement, len will default to 100:

```
lister(my_list)
```

Some programming languages (e.g., Ada and Python) support keyword parameters. In a keyword parameter, the actual parameter specifies the name of its associated formal parameter, as in the following statement:

```
lister(list => my_list, len => 50)
```

The advantage of keyword parameters is that they eliminate the possibility of making mistakes in the association of actual parameters with formal parameters. This property is particularly useful

when there are more than a few parameters.

Ruby does not support keyword parameters, but there is a way to achieve the same benefit with hashes. A hash literal has an appearance that is similar to keyword parameters. For example, if a hash literal is passed as a parameter to the find method, the hash literal would appear as follows:

```
find(age, {'first' => 'Davy', 'last' => 'Jones'})
```

Whenever such a hash literal is passed as an actual parameter and it follows all normal scalar parameters and precedes all array and block parameters, the braces can be omitted. So, in the preceding example, the braces are unnecessary.

Ruby includes a category of objects that appears in no other common programming languages: symbols. Symbols are created by preceding an unquoted string with a colon (:).² A symbol made from a variable name can be thought of as that variable's name. Such a symbol does not refer to the value of the variable, nor is it related to a particular instance of a variable—so symbols are context independent. All symbols are instances of the Symbol class. Symbols can be used to specify parameters in method calls and the keys of elements of hash literals. It has become a Ruby idiom, and even a convention in Rails, to use symbols, rather than literal strings, for the keys in hash literals when they are used as parameters. The following method call is illustrative:

```
find(age, :first => 'Davy', :last => 'Jones')
```

Following is a method that computes the median of a given array of numbers:

```
# median - a method
# Parameter: An array of numbers
# Return value: The median of the parameter array
#
def median(list)

  # Sort the array
  list2 = list.sort

  # Get the length of the array
  len = list2.length

  # Compute the median
  if(len % 2 == 1)  # length is odd
    return list2[len / 2]
  else              # length is even
    return (list2[len / 2] + list2[len / 2 - 1]) / 2
  end

end # end of the median method
```

8.8 Classes

Classes in Ruby are like those of other object-oriented programming languages, at least in purpose. A class defines the template for a category of objects, of which any number can be created. An object has a state, which is maintained in its collection of instance variables, and a behavior, which is defined by its methods. An object can also have constants and a constructor.

The Basics of Classes

The methods and variables of a class are defined in the syntactic container that has the following form:

```
class class_name
```

```
...
```

```
end
```

Class names, like constant names, must begin with uppercase letters.

Instance variables are used to store the state of an object. They are defined in the class definition, and every object of the class gets its own copy of the instance variables. The name of an instance variable must begin with an at sign (@), which distinguishes instance variables from other variables.

A class can have a single constructor, which in Ruby is a method with the name `initialize`, which is used to initialize instance variables to values. A constructor can take any number of parameters, which are treated as local variables; therefore, their names begin with lowercase letters or underscores. The parameters are given after the call to `new`.

Following is an example of a class, named `Stack2_class`, that defines a stacklike data structure implemented in an array. The difference between this structure and a stack is that both the top element and the element that is second from the top are accessible. The latter element is fetched with the `top2` method. Here is the class:

```

# Stack2_class.rb - a class to implement a stacklike
#                      structure in an array
class Stack2_class

# Constructor - parameter is the size of the stack - default is 100
def initialize(len = 100)
    @stack_ref = Array.new(len)
    @max_len = len
    @top_index = -1
end

# push method
def push(number)
    if @top_index == @max_len
        puts "Error in push - stack is full"
    else
        @top_index += 1
        @stack_ref[@top_index] = number
    end
end

# pop method
def pop()
    if @top_index == -1
        puts "Error in pop - stack is empty"
    else
        @top_index -= 1
    end
end

```

Following is simple code to illustrate the use of the Stack2_class class:

```

# Test code for Stack2_class
mystack = Stack2_class.new(50)
mystack.push(42)
mystack.push(29)
puts "Top element is (should be 29): #{mystack.top}"
puts "Second from the top is (should be 42): #{mystack.top2}"
mystack.pop
mystack.pop
mystack.pop # Produces an error message - empty stack

```

Classes in Ruby are dynamic in the sense that members can be added at any time, simply by including additional class definitions that specify the new members. Methods can also be removed from a class, by providing another class definition in which the method to be removed is sent to the method `remove_method` as a parameter. The dynamic classes of Ruby are another example of a language designer trading readability (and, as a consequence, reliability) for flexibility. Allowing dynamic changes to classes clearly adds flexibility to the language, but harms readability. To determine the current definition of a class, one must find and consider all of its definitions in the program.

Access Control

In a clear departure from the other common programming languages, access control in Ruby is different for access to data than it is for access to methods. All instance data has private access by default, and that access cannot be changed. If external access to an instance variable is required, access methods must be defined. For example, consider the following skeletal class definition:

```
class My_class

# Constructor
def initialize
  @one = 1
  @two = 2
end

# A getter for @one
def one
  @one
end

# A setter for @one
def one=(my_one)
  @one = my_one
end

end # of class My_class
```

The equals sign (=) attached to the name of the setter method means that the method is assignable. So, all setter methods have equals signs attached to their names. The body of the one method illustrates the Ruby design whereby methods return the value of the last expression evaluated when there is no return statement. In this case, the value of @one is returned. When an instance variable that has a getter or setter is referenced outside the class, the at sign (@) part of the name is not included. The following code that uses My_class is illustrative:

```
mc = My_class.new
puts "The value of one is #{mc.one}"
```

Because getter and setter methods are frequently needed, Ruby provides shortcuts for both. If one wants a class to have getter methods for two instance variables, @one and @two, those getters can be specified with the single statement in the class as follows:

```
attr_reader :one, :two
```

Note that attr_reader is actually a method call, using the symbols :one and :two as the actual parameters.

The function that similarly creates setters is called attr_writer. This function has the same

parameter profile as attr_reader.

The functions for creating getter and setter methods are so named because they provide the protocol for object members of the class, which are called attributes in Ruby. So, the attributes of a class constitute the data interface (the public data) to objects of the class.

The three levels of access control for methods are defined as follows: “Public access” means that the method can be called by any code. “Protected access” means that only objects of the defining class and its subclasses may call the method. “Private access” means that the method cannot be called with an explicit receiver object. Because the default receiver object is self, a private method can be called only in the context of the current object. So, no code can ever call the private methods of another object. Note that private access in Ruby is quite different from private access in other programming languages, such as C++, Java, and C#.

Access control for methods in Ruby is dynamic, so access violations are detected only during execution. The default method access is public, but it can also be protected or private. There are two ways to specify the access control, both of which use functions with the same names as the access levels: private, protected, and public. One way is to call the appropriate function without parameters. This resets the default access for all subsequent defined methods in the class, until a call to a different access control method appears. The following class illustrates the first way:

```
class My_class
  def meth1
    ...
  end
  ...
  private
  def meth7
    ...
  end
  ...
  protected
  def meth11
    ...
  end
  ...
end # of class My_class
```

The alternative is to call the access control functions with the names of the specific methods as parameters. For example, the following is semantically equivalent to the previous class definition:

```

class My_class
  def meth1
  ...
end
...
def meth7
...
end
...
def meth11
...
end
...
private :meth7, ...
protected :meth11, ...
end # of class My_class

```

The default access control for constructors is private. Class variables are private to the class and its instances. That privacy cannot be changed. Also, unlike global and instance variables, class variables must be initialized before they are used.

Inheritance

Subclasses are defined in Ruby with the left angle bracket (<):

```
class My_Subclass < Base_class
```

One distinctive feature of Ruby's method access controls is that they can be changed in a subclass simply by calling the access control functions. This means that two subclasses of a base class can be defined so that objects of one of the subclasses can access a method defined in the base class, but objects of the other subclass cannot. Also, it allows one to change the access of a publically accessible method in the base class to a privately accessible method in the subclass. Such a subclass obviously cannot be a subtype.

Ruby modules provide a naming encapsulation that is often used to define libraries of methods. Perhaps the most interesting aspect of modules, however, is that their methods can be accessed directly from classes. Access to a module in a class is specified with an include statement, such as the following:

```
include Math
```

The effect of including a module is that the class gains a pointer to the module and effectively inherits the functions defined in the module. In fact, when a module is included in a class, the

module becomes a proxy superclass of the class. Such a module is called a mixin, because its functions get mixed into the methods defined in the class. Mixins provide a way to include the functionality of a module in any class that needs it—and, of course, the class still has a normal superclass from which it inherits members. So, mixins provide the benefits of multiple inheritance, without the naming collisions that could occur if modules did not require module names on their functions.

8.9 Blocks and Iterators

A block is a sequence of code, delimited by either braces or the do and end reserved words. Blocks can be used with specially written methods to create many useful constructs, including simple iterators for arrays and hashes. This construct consists of a method call followed by a block. In this section, a few of the built-in iterator methods that are designed to use blocks are discussed.

The times iterator method provides a way to build simple counting loops. Typically, times is sent to a number object, which repeats the attached block that number of times. Consider the following example:

```
>> 4.times {puts "Hey!"}
```

```
Hey!
```

```
Hey!
```

```
Hey!
```

```
Hey!
```

```
=> 4
```

In this example, the times method repeatedly executes the block in a different approach to subprogram control. (A block is clearly a form of a subprogram.)

The most common iterator is each, which is often used to go through arrays and apply a block to each element. For this purpose, it is convenient to allow blocks to have parameters, which, if present, appear at the beginning of the block, delimited by vertical bars (|). The following example, which uses a block parameter, illustrates the use of each:

```
>> list = [2, 4, 6, 8]
=> [2, 4, 6, 8]
>> list.each {|value| puts value}
2
4
6
8
=> [2, 4, 6, 8]
```

The each iterator works equally well on array literals, as in the following interactions:

```
>> ["Joe", "Jo", "Joanne"].each {|name| puts name}
Joe
Jo
Joanne
=> ["Joe", "Jo", "Joanne"]
```

If each is called on a hash, two block parameters must be included, one for the key and one for the value:

```
>> high_temps = {"Mon"=>72, "Tue"=>84, "Wed"=>80}
=> {"Wed"=>80, "Mon"=>72, "Tue"=>84}
>> high_temps.each
      {|day, temp| puts "The high on #{day} was #{temp}"}
The high on Wed was 80
The high on Mon was 72
The high on Tue was 84
=> {"Wed"=>80, "Mon"=>72, "Tue"=>84}
```

The upto iterator method is used like times, except that the last value of the counter is given as a parameter:

```
>> 5.upto(8) {|value| puts value}
5
6
7
8
=> 5
```

The step iterator method takes a terminal value and a step size as parameters and generates the values from that of the object to which it is sent and the terminal value:

```
>> 0.step(6, 2) {|value| puts value}
0
2
4
6
=> 0
```

Like each, the collect iterator method takes the elements from an array, one at a time, and puts the values generated by the given block into a new array:

```
>> list = [5, 10, 15, 20]
=> [5, 10, 15, 20]
>> list.collect {|value| value = value - 5}
=> [0, 5, 10, 15]
>> list
=> [5, 10, 15, 20]
>> list.collect! {|value| value = value - 5}
=> [0, 5, 10, 15]
>> list
=> [0, 5, 10, 15]
```

As can be seen from this example, the mutator version of collect is probably more often useful than the nonmutator version, which does not save its result.

Now we consider user-defined methods and blocks. There must be some statement in the method that “calls” the block. This statement is yield. The yield statement is similar to a method call, except that there is no receiver object and the call is a request to execute the block attached to the method call, rather than a call to a method. If the block has parameters, they are specified in parentheses in the yield statement. The value returned by a block is that of the last expression evaluated in the block. A method can include any number of yield statements, so it can cause the block to be “called” any number of times. It is this process that is used to implement the iterators illustrated earlier in this section.

When a block is used in a call to a method, part of the effect of the call is provided by the code in the method and part is provided by the block. This separation of functionality allows a method to have different effects on different calls, with the different effects provided by the block attached to the call. The following example is illustrative:

```
>> def get_name
>>   puts "Your name:"
>>   name = gets
>>   yield(name)
>> end
=> nil
>> get_name {|name| puts "Hello, " + name}
Your name:
Freddie
Hello, Freddie
=> nil
```

8.10 Pattern Matching

The Basics of Pattern Matching

In Ruby, the pattern-matching operation is specified with the matching operators `=~`, for positive matches, and `!~`, for negative matches. Patterns are placed between slashes (`/`). For example, in the following interactions the right operand pattern is matched against the left operand string:

```
>> street = "Hammel"
=> "Hammel"
>> street =~ /mm/
=> 2
```

The result of evaluating a pattern-matching expression is the position in the string where the pattern matched.

The split method is frequently used in string processing. The method uses its parameter, which is a pattern, to determine how to split the string object to which it is sent into substrings. For example, the interactions

```
>> str = "Jake used to be a small child, but now is not."
=> "Jake used to be a small child, but now is not."
>> words = str.split(/[,]\s*/)
=> ["Jake", "used", "to", "be", "a", "small", "child",
"but", "now", "is", "not"]
```

puts the words from str into the words array, where the words in str are defined to be terminated with either a space, a period, or a comma, any of which could be followed by more white-space characters.

The example program that follows illustrates a simple use of pattern matching and hashes. The program reads lines of text in which the words are separated by white space and some common kinds of punctuation, such as commas, periods, semicolons, and so forth. The objective of the program is to produce a frequency table of the words found in the input. A hash is an ideal way to build the word-frequency table. The keys can be the words, and the values can be the number of times they have appeared. The split method provides a convenient way to split each line of the input file into its component words. For each word, the program uses `has_key?` on the hash to determine whether the word has occurred before. If so, its count is incremented; if not, the word is entered into the hash with a count of 1. Here is the code:

```

# Input: Text from the keyboard. All words in the input are
# separated by white space or punctuation, possibly followed
# by white space, where the punctuation can be a comma, a
# semicolon, a question mark, an exclamation point, a period,
# or a colon.
# Output: A list of all unique words in the input, in alphabetical
# order, along with their frequencies of occurrence

freq = Hash.new
line_words = Array.new

# Main loop to get and process lines of input text
while line = gets

    # Split the line into words
    line_words = line.chomp.split( /[ \.,;!:?\]\s*/)

    # Loop to count the words (either increment or initialize to 1)
    for word in line_words
        if freq.has_key?(word) then
            freq[word] = freq[word] + 1
        else
            freq[word] = 1
        end
    end
end
# Display the words and their frequencies
puts "\n Word \t\t Frequency \n\n"
for word in freq.keys.sort
    puts " #{word} \t\t #{freq[word]}"
end

```

Notice that the two normally special characters, . (period) and ? (question mark), are not backslashed in the pattern for split in this program. This is because the normally special characters for patterns (metacharacters) are not special in character classes.

Remembering Matches

The part of the string that matched a part of the pattern can be saved in an implicit variable for later use. The part of the pattern whose match you want to save is placed in parentheses. The substring that matched the first parenthesized part of the pattern is saved in \$1, the second in \$2, and so forth. The following interactions show how this is done:

```
>> str = "4 July 1776"
=> "4 July 1776"
>> str =~ /(\d+) (\w+) (\d+)/
=> 0
>> puts "${2} ${1}, ${3}"
=> July 4, 1776
```

In some situations, it is convenient to be able to reference the part of the string that preceded the match, the part that matched, or the part that followed the match. These three strings are available after a match through the implicit variables \$`, \$&,amp;, and \$', respectively.

Substitutions

Sometimes the substring of a string that matched a pattern must be replaced by another string. Ruby's String class has four methods designed to do exactly that. The most basic of these, the substitute method, sub, takes two parameters: a pattern and a string (or an expression that evaluates to a string value). The sub method matches the pattern against the string object to which it is sent. If sub finds a match, the matched substring is replaced by its second parameter, as in the following interactions:

```
>> str = "The old car is great, but old"
=> "The old car is great, but old"
>> str.sub(/old/, "new")
=> "The new car is great, but old"
```

The gsub method is similar to sub, except that it finds all substring matches and replaces all of them with its second parameter:

```
>> str = "The old car is great, but old"
=> "The old car is great, but old"
>> str.gsub(/old/, "new")
=> "The new car is great, but new"
>> str
=> "The old car is great, but old"
```

Notice from the last line that gsub does not alter the string object on which it is called. The same is true for sub. However, sub and gsub have mutator versions, named sub! and gsub!. The following interactions illustrate how gsub! works:

```
>> str = "The old car is great, but old"
=> "The old car is great, but old"
>> str.gsub!(/old/, "new")
=> "The new car is great, but new"
>> str
=> "The new car is great, but new"
```

The `i` modifier, which tells the pattern matcher to ignore the case of letters, can also be used with the substitute method by attaching it to the right end of the pattern, as shown in the following code:

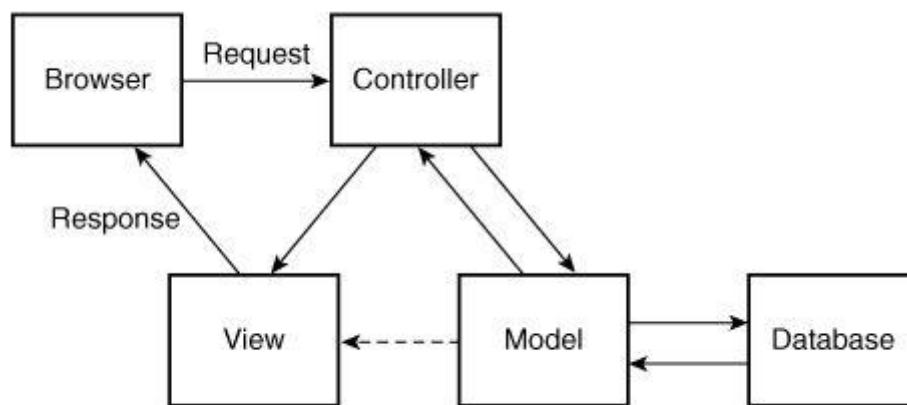
```
>> str = "Is it Rose, rose, or ROSE?"
=> "Is it Rose, rose, or ROSE?"
>> str.gsub(/rose/i, "rose")
=> "Is it rose, rose, or rose?"
```

8.11 Overview of Rails

Rails are a software development framework for Web-based applications—in particular, those that access databases. A framework is a system in which much of the more-or-less standard software parts are furnished by the framework, so they need not be written by the applications developer. Those parts are often skeletal classes, methods, or markup documents, but can also be complete utility methods. Rails was developed by David Heinemeier Hansson in the early 2000s and was released to the public in July 2004. Since then, it has rapidly gained widespread interest and usage.

Rails, like some other Web development frameworks, such as Tapestry and Struts, is based on the Model–View–Controller (MVC) architecture for applications.

Figure below shows the components and actions of a request and response in a Rails application that uses a database.



A request and response in a Rails application

The view part of a Rails application generates the user interface. Both data in the model and results of processing are made available to the user through view documents. Also, some view documents

provide the interface for administrators to add, modify, or delete data. View documents are markup documents that can have embedded Ruby code, which is interpreted on the server before the documents are sent to a browser, much like what happens with requested PHP documents.

The controller part of a Rails application, which is implemented as one or more Ruby classes, controls the interactions among the data model, the user, and the view. The controller receives user input, interacts with the model, and provides views of data and processing results back to the user. The developer must design and build the actions that are required by the application, implemented as methods in the controller classes.

The model part of a Rails application maintains the state of the application, whether that state is internal and alive only during execution or is a permanent external database. The developer must design and build a model of the application's domain. The design of the model often includes a database that stores the data of the model. For example, if the application is an online bookstore, the model might include an inventory of books and a catalog of all books, among other things, that can be ordered through the store. The model also can include constraints on the data to be entered into the database.

So, what does Rails do for the developer of a Rails application? It does quite a lot, actually. Rails provides skeletal controller classes. It also implicitly connects the methods of a controller with the corresponding view documents. In addition, it provides the basic interface to a working database, as well as an empty version of the database itself. One of its most important contributions is a collection of conventions that implicitly connect the model, view, and controller. For example, the controller can fetch user form data and place it in instance variables, which are implicitly available to the Ruby code in view documents. Rails also provides partial view documents, known as layouts, and a simple way to include style sheets for documents. In addition, Rails provides a simple development and test environment, including Web servers—and Rails is free.

A Rails application is a program that provides a response when a client browser connects to a Rails-driven Web site. Because Rails uses an MVC architecture, building a Rails application consists of designing and building the three components of an MVC system. Rails offers a great deal of assistance in constructing an application, as will be evidenced in the example applications in this chapter.

There are two fundamental principles that guided the development of Rails, and it is valuable to be aware of them when learning and using Rails. The first principle has the acronym DRY, which stands for Do not Repeat Yourself. In Rails, DRY means that every element of information appears just once in the system. This minimizes the memory required by the system. In addition, changes to the system are highly localized, making them both easier and less error prone. The second principle is named convention over configuration. Web applications with JSP require elaborate and complicated XML configuration files to specify their structure. In Rails, the structure of an application is dictated by the MVC architecture. The connections between the different parts are established and maintained by convention, rather than being specified in a configuration document. For example, the names of database tables and their associated controller classes are intimately related by convention.

Rails is a product of a software development paradigm called agile development.³ Some of this paradigm is related to the human interactions among development team members and between the team and the customer. However, part of it is the focus on the quick development of working software, rather than the creation of elaborate documentation and then software. Agile development is an incremental approach to development that is facilitated by adherence to the principles used in creating Rails.

Rails can be, and often is, used to develop Ajax-enabled Web applications. Rails uses a JavaScript library named Prototype to support Ajax and interactions with the JavaScript model of the document being displayed by the browser. Rails also provides other support for developing Ajax, including producing visual effects.

Rails differs from the other frameworks discussed in this book—Flash, Net-Beans, and Visual Studio—in that it does not use a graphical user interface (GUI). Rather, Rails is a command-line-oriented system. Commands are issued by typing them at a prompt in a DOS-like or UNIX-like command window, rather than by clicking icons on a GUI.

Some of the innovations of Rails are described through examples presented later in the chapter, among the most interesting of which are the basic database operations furnished by Rails for a new database application and the use of migrations to manage version control of databases.

8.12 Document Requests

Rails is a Web application development framework. A software development framework is often constructed as a library of components that provide commonly needed services to an application in a particular application area.

Before one can use Rails, the system must be downloaded and installed on one's computer. For Windows users, one simple way to do this is to download the complete software system named Instant Rails from <http://instantrails.ruby-forge.org/wiki/wiki.pl>. Instant Rails, which was developed by Curt Hibbs, includes Ruby, Rails, MySQL, several Web servers, and everything else needed to use these technologies together. Installing Instant Rails is quick and easy.

Instant Rails is a self-contained system. It does not reside in the global Windows environment, so interactions with it must be done through a special command window.

The Leopard operating system (Mac OS X Version 10.5) for Macintosh computers was released in October 2007. This system includes Ruby and Rails, so no downloading or installation is required to use Rails on these Macs running that operating system.

For UNIX systems, Rails is available from the following site:

<http://www.RubyonRails.org/down>

The Rails examples in this chapter were developed with Rails 2.3.4.

This section describes how to build a Hello, World application in Rails. The purpose of such an exercise is to demonstrate the directory structure of the simplest possible Rails application, showing what files must be created and where they must reside in the directory structure.

If Instant Rails is to be used, first it must be started. This is done by clicking the thick I, which is the icon of the Instant Rails application, in the directory in which Instant Rails was installed. A small window then opens, as shown in Figure.



The Instant Rails application window

A click on the black I in the upper-left part of this window produces a small menu. Selecting the Rails Applications entry in this menu opens another menu, whereupon selecting the Open Ruby Console Window entry opens a command-line window in the rails_apps subdirectory of the directory in which Instant Rails was installed. This in turn opens a Rails command window, which is similar in appearance to a Windows Command Prompt window. In using Instant Rails, Rails commands cannot be given in a normal command window—only in a Rails command window.

If you are running Windows and the IIS server, you may need to stop that server before using Rails. Also, if MySQL is running, it, too, must be stopped, because InstantRails will start its own MySQL.

On a UNIX or Macintosh system, Rails is run from the normal command line in the directory in which Rails was installed.

Users usually create a new subdirectory of rails_apps for their Rails applications. We created a subdirectory named examples for the example applications of this chapter.

Next, we move to the examples directory and create a new Rails application named greet with the following command:

```
>rails greet
```

Rails responds by creating more than 45 files in more than 30 directories. This is part of the framework that supports a Rails application. Directly under the specific application directory—in this case, greet—11 subdirectories are created, the most interesting of which at this point is app. The app directory has four subdirectories: models, views, and controllers, which correspond

directly to the MVC architecture of a Rails application—and helpers. The helpers subdirectory contains Rails-provided methods that aid in constructing applications. Most of the user code to support an application will reside in models, views, or controllers, or in subdirectories of those directories.

One of the directories created by the rails command is script, which has several important Ruby scripts that perform services. One such script, generate, is used to create part of an application controller. This script creates a file containing a class in the controllers directory, and also a subdirectory of the views directory where views documents will be stored. For our application, we pass two parameters to generate, the first of which is controller, which indicates that we want the controller class built. The second parameter is the name we chose for the controller. An important part of how Rails works is its focused use of names. Our first example of this feature is the name of the controller, which will also be part of the file name of the controller class and part of the name of the controller class itself. In addition, it will be the name of the subdirectory of the views directory and a part of the URL of the application. For our example, the following command is given in the greet directory to create the controller:

```
>ruby script/generate controller say
```

With this command, we have chosen the name say for the controller for our application. The response produced by the execution of the command is as follows:

```
exists  app/controllers/
exists  app/helpers/
create  app/views/say
exists  test/functional/
create  app/controllers/say_controller.rb
create  test/functional/say_controller_test.rb
create  app/helpers/say_helper.rb
```

The exists lines in this response indicate files and directories that Rails found to already exist. The create lines show the newly created directories and files. There are now two files in the controllers directory—application.rb and say_controller.rb—which contain the ApplicationController and SayController classes, respectively. The SayController class is a subclass of the ApplicationController class. As the parent class, ApplicationController provides the default behavior for SayController, the controller class of the application class. There may be other controllers and their corresponding controller classes in an application. Such classes are subclasses of ApplicationController, which was built by the initial rails command. The following is a listing of say_controller.rb:

```
class SayController < ApplicationController
end
```

Note the occurrence of say in both the name of the controller file and the name of the controller

class. This is another example of the use of convention in Rails.

The class SayController is an empty class, other than what it inherits from ApplicationController, which is in the file application.rb. The class ApplicationController is a subclass of ActionController, which defines the basic functionality of a controller. Note that SayController produces, at least indirectly, the response to requests, so a method must be added to it. The method does not need to actually do anything, other than indicate a document that will describe the response.⁴ The mere existence of the method specifies, by its name, the response document. So, the action will be nothing more than an empty method definition whose name will be the same as that of the response document in the say subdirectory of views. With the empty method, which is called an action method, the controller now has the following appearance:

```
class SayController < ApplicationController
  def hello
  end
end
```

Web sites, or applications, are specified in requests from browsers with URLs. Rails applications are no different. When Rails receives the URL of a request, it maps that URL to a specific controller name and action method. In simple situations, the mapping is trivial: the first domain following the hostname is interpreted as a controller name, and the next domain is interpreted as the name of an action method. There is no need to specify the application name, because, as we shall soon see, each application is served by its own server.

The host for our examples will be the machine on which the applications are resident. The default port for the server (chosen by Rails) is 3000, so the host name will be localhost:3000. Thus, for the greet example, the request URL is as follows:

<http://localhost:3000/say/hello>

(Now it should be obvious why the base document is named say.)

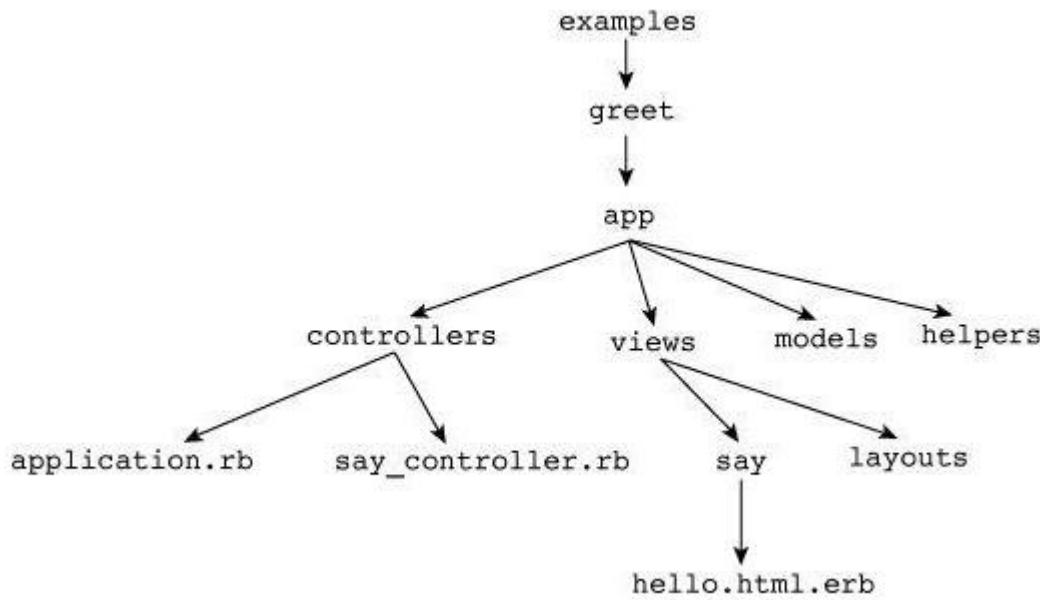
Next, we need to build the response document, or view file, which will be a simple XHTML file to produce the greeting. The view document is often called a template. The following is the template for the greet application:

```
<!-- hello.html.erb - the template for the greet application
-->
<html>
  <head>
    <title> greet </title>
  </head>
  <body>
    <h1> Hello from Rails! </h1>
  </body>
</html>
```

Note that, for simplicity's sake, we have included neither an XML nor a DOCTYPE declaration. The extension on this file's name is .html.erb because the file stores an HTML document, but it may include embedded Ruby code to be interpreted by the Ruby interpreter, ERb (an acronym for Embedded Ruby), before the template is returned to the requesting browser.

The template file for our application resides in the say subdirectory of the views subdirectory of the app subdirectory of the greet directory.

The structure of the examples directory is shown in Figure.



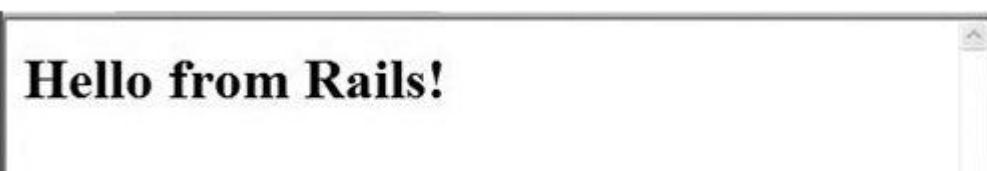
Directory structure for the greet application

Before the application can be tested, a Rails Web server must be started. A server is started with the server script from the script directory. The default server is Mongrel, but the Apache and WEBrick servers are also available within Rails. Because it is the default Rails server, Mongrel can be started with the following command at the application prompt:

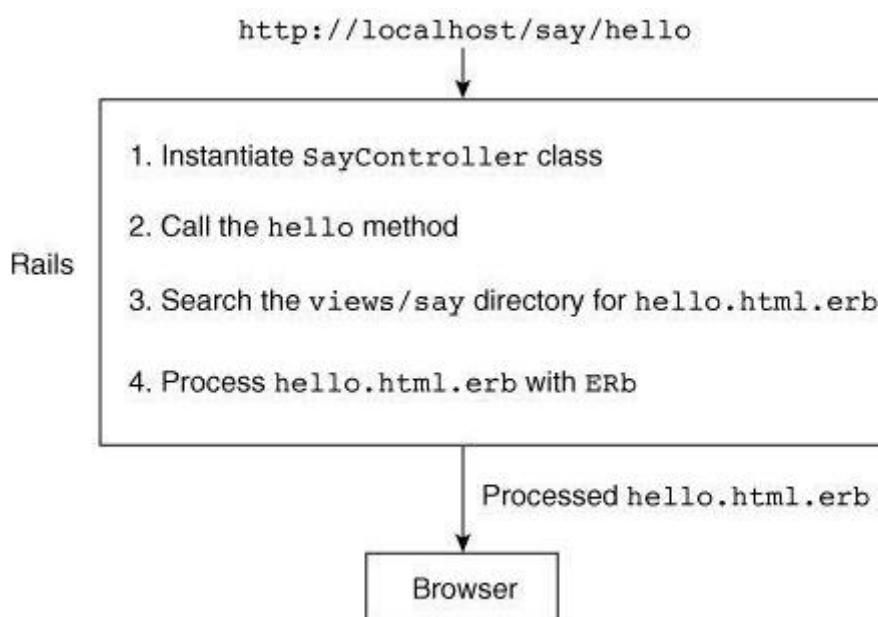
```
>ruby script/server
```

Note that the server is started by a command in the directory of the particular application—in our example, greet. This implies that no other application can be served by this server.

Figure below shows the output of the greet application when it is addressed by a browser.



The following summarizes how Rails reacts to a request for a static document: First, the name of the controller is extracted from the URL. (It follows the hostname.) Next, an instance of the controller class (found in the app/controllers subdirectory)—in our example, SayController—is created. The name of the action is then extracted from the URL—in our example, hello. Then Rails searches for a template with the same name as the action method in the subdirectory with the same name as the controller in the app/views directory. Next, the template file is given to ERb to interpret any Ruby code that is embedded in the template. In the case of hello.html.erb, there is no embedded Ruby code, so this step has no effect. Finally, the template file is returned to the requesting browser, which displays it. The activities of Rails in response to a simple request are shown in Figure.



The default method of a controller class is `index`. If the URL `http://localhost:3000/say/` was requested, Rails would search for an `index` method, which, in this example, does not exist. The result would be the following message on the display: Unknown action - no action responded to

index.

Dynamic Documents

Dynamic documents can be constructed in Rails by embedding Ruby code in a template file. This approach is similar to some other approaches we have discussed—in particular, PHP, ASP.NET, and JSP.

As an example of a dynamic document, we modify the greet application to display the current date and time on the server, including the number of seconds since midnight (just so that some computation will be included). This modification will illustrate how Ruby code that is embedded in a template file can access instance variables that are created and assigned values in an action method of a controller.

Ruby code is embedded in a template file by placing it between the `<%` and `%>` markers. For example, we could insert the following in a template:

```
<% 3.times do %>  
I LOVE YOU MORE! <br />  
<% end %>
```

If the Ruby code produces a result and the result is to be inserted into the template document, an equals sign (`=`) is attached to the opening marker. For example, the element

```
<p> The number of seconds in a day is: <%= 60 * 60 * 24 %>  
</p>
```

produces

```
<p> The number of seconds in a day is: 86400 </p>
```

after interpretation of the Ruby code.

The date can be obtained by calling Ruby's `Time.now` method, which returns the current day of the week, month, day of the month, time, time zone, and year, as a string. So, we can put the date in the response template with

```
<p> It is now <%= Time.now %> </p>
```

The value returned by `Time.now` can be parsed with the methods of the `Time` class. For example, the `hour` method returns the hour of the day, the `min` method returns the minutes of the hour, and the `sec` method returns the seconds of the minute. These methods can be used to compute the number of seconds since midnight. Putting it all together results in the following template code:

```
It is now <%= t = Time.now %> <br />
Number of seconds since midnight:
<%= t.hour * 3600 + t.min * 60 + t.sec %>
```

It would be better to place the Ruby code for the time computation in the controller, because that would separate the program code from the markup. The modified SayController class is as follows:

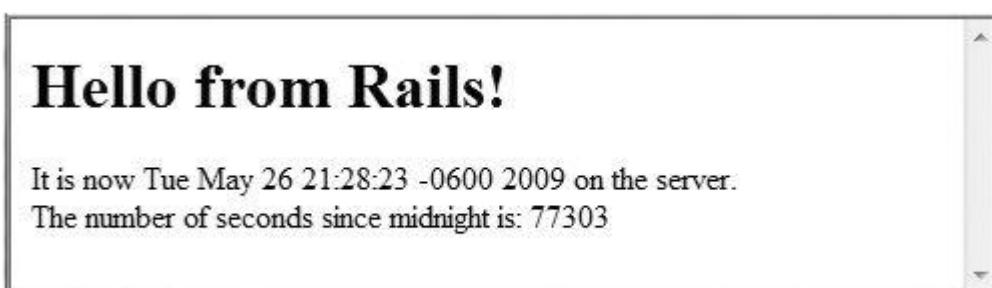
```
class SayController < ApplicationController
  def hello
    @t = Time.now
    @tsec = @t.hour * 3600 + @t.min * 60 + @t.sec
  end
end
```

The response template now needs to be able to access the instance variables in the SayController class. Rails makes this process trivial, for all instance variables in the controller class are visible to the template. The template code for displaying the time and number of seconds since midnight is as follows:

It is now <%= @t %>

Number of seconds since midnight: <%= @tsec %>

Figure shows the display of the modified greet application.



Some care must be taken in choosing variable names. Rails includes a large number of reserved words, which, if used as variable names, cause problems that are difficult to diagnose. The list of Rails reserved words is given at <http://wiki.RubyonRails.org/rails/show/ReservedWords>.

8.13 Rails Applications with Databases

A significant and characteristic part of Rails is its approach to connecting object-oriented software with a relational database. These two things are not particularly amenable to marriage, so the connection is not a natural one. Rails uses an object-relational mapping (ORM) approach to relate the parts of a relational database to object-oriented constructs. Each relational database table is implicitly mapped to a class. For example, if the database has a table named employees, the Rails

application program that uses employees will have a class named Employee. Rows of the employees table will have corresponding objects of the Employee class, which will have methods to get and set the various state variables, which are Ruby attributes of objects of the class. In sum, an ORM maps tables to classes, rows to objects, and columns to the fields of the objects. Furthermore, the Employee class will have methods for performing table-level operations, such as finding an object with a certain attribute value. The key aspect of the ORM in Rails is that it is implicit: The classes, objects, methods, and attributes that represent a database in Ruby are automatically built by Rails.

For this example, we create a new application named cars in the examples directory with the following command:

```
>rails -d mysql cars
```

The d flag followed by mysql appears in the rails command in order to tell Rails that this application will use a MySQL database.