Mark Cyzyk. Book Review: Clinton Wong's *Web Client Programming with Perl*, in *Telecommunications Electronic Reviews*, v4 (8), September 1997.

ter

telecommunications electronic reviews

Telecommunications Electronic Reviews (TER) is a publication of the Library and Information Technology Association.

REVIEW OF: Clinton Wong. Web Client Programming with Perl. Sebastopol, CA: O'Reilly & Associates, 1997.

by Mark Cyzyk

The purpose of this book is to introduce UNIX-based Webmasters to simple Web client programming using Perl (Practical Extraction and Report Language). It can, however, be profitably read by anyone interested in how Web clients work.

The first third of the volume concentrates on what Web clients are and how they do what they do. The second chapter, "Demystifying the Browser," examines the anatomy of a Web client and the basic structure of a client/server transaction. The various HTTP (Hypertext Transfer Protocol) headers that are passed from client to server and back again are illustrated. To illustrate the text-based nature of HTTP, Wong offers an interesting example of a manual HTTP transaction. He shows how one can simply Telnet into a Web server and manually issue HTTP commands resulting in output from the

server. This powerful example of a simple HTTP transaction stays firmly in the reader's mind during the rest of the book.

The third chapter, "Client Request Methods," offers a relatively in-depth discussion of HTTP itself. Although there are fully ten methods that a client can use to make a request of a server, Wong spends time only on the seven most important: GET, HEAD, POST, PUT, DELETE, TRACE, and OPTIONS. These are the most useful request methods for the budding client programmer to understand. All CGI (Common Gateway Interface) programmers are familiar with the GET and POST methods, but Wong's discussion of the other methods is fascinating to those like myself who, when confronted with the automatic upload features of recent versions of Netscape and other Web-based applications, wonder "How did they do that?" (Hint: they used the PUT method.) The chapter ends with a solid discussion of server response codes, e.g., 403 Forbidden, and what they mean, as well as a comprehensive list and select discussion of HTTP headers.

For the client programmer who revels in the bits and bytes, chapter four, "The Socket Library," covers network socket programming over a TCP/IP (Transmission Control Protocol/Internet Protocol) network, specifically in a UNIX environment. Though important for a comprehensive understanding of the client/server process over TCP/IP, as Wong himself states, a deep understanding of socket programming is not needed to create Web-based client programs. The reason for this is that programmers can avail themselves of the Library of World Wide Web (WWW) modules for Perl (LWP) that are widely available from most Perl archives.

Like any Perl library, the LWP modules exist to greatly simplify the programming process. The heart of the book is to be found in chapters five and six, "The LWP Library" and "Example LWP Programs," where the LWP Library is discussed and its uses are illustrated. LWP is a library of modules for Perl 5 that allows the programmer to accomplish complex tasks (such as socket programming or Uniform Resource Locator parsing) in a simple and elegant manner.

The library as a whole is comprised of eight discrete modules of which Wong addresses only the four most useful: the LWP, HTML (Hypertext Markup Language), HTTP, and URI (Uniform Resource Identification) modules. Essentially, the LWP module handles socket communications between client and server; the HTML module handles the parsing of HTML documents; the HTTP module handles HTTP requests and responses; and the URI module handles the escaping of URIs and the relative-to-absolute translation of URLs. Within each module are to be found several classes (LWP is object-oriented); Wong discusses the most important for each of the four modules.

Wong offers three types of example Web clients--simple clients, periodic clients, and recursive clients--to illustrate the many uses of the LWP library. Wong's simple client calls a URL with a custom User Agent header designating the name of the client program, thereby identifying itself to the server. His periodic client--more complex than the simple client--periodically connects to the Federal Express Web site, uses a tracking number to query its database, and waits to see if a package has been delivered. The program runs until it receives notification of delivery. Such a periodic client could form the basis of any Web-based function that needs to be performed regularly, for example, a stock ticker checker. Wong's recursive client--by far the most complex--checks all the HTML pages at a specified Web site for malfunctioning hyperlinks. A list of these links is then printed out.

For those running X Windows, the final chapter, "Graphical Examples with Perl/Tk," illustrates how one can program a graphical user interface (GUI) for Perl text-based programs. In this manner, the

programmer can actually program a custom GUI Web browser.

As are most of the O'Reilly publications, this book is well- prepared, useful, and a pleasure to read, using the chatty style that O'Reilly authors seem to prefer. Examples throughout are relevant and clearly presented. The book as a whole proceeds in a satisfying manner, its pace easing the reader into ever deeper programming concepts and examples. For the Webmaster hoping to create custom, text-based Web clients for use in Web site administration, or for those just interested in learning more about how the hypertext transfer protocol functions, this book is recommended.

Mark Cyzyk (mcyzyk@towson.edu) is Head of Information Technology at Albert S. Cook Library, Towson State University, in Maryland.

About TER

Technology Electronic Reviews (TER) is an irregular electronic serial publication of the Library and Information Technology Association, a division of the American Library Association, 50 E. Huron St., Chicago, IL 60611. The primary function of TER is to provide reviews of and pointers to a variety of print and electronic resources about information technology. Resources include books, articles, serials, discussion lists, training materials, bibliographies, and other items of interest to librarians and information technology professionals. The topics covered may include, but are not limited to, networking technologies and standards; hardware and software; operating systems; databases; specific programming languages; management tools and utilities; technical project management; training and personnel issues; library perspectives; and research and development.

Opinions expressed in this publication are those of the writers and do not necessarily represent the viewpoints of LITA, ALA, or organizations involved in the storage and/or distribution of the publication.

TER is distributed electronically via Internet. There is no subscription fee. Currently it is available via World Wide Web (http://www.lita.org/ter/) and new-issue announcements are posted on the LITA-L

electronic discussion list. To subscribe, send an email message to listproc@ala1.ala.org that says: subscribe LITA-L First-Name Last-Name. Other distribution arrangements may be made in the future.