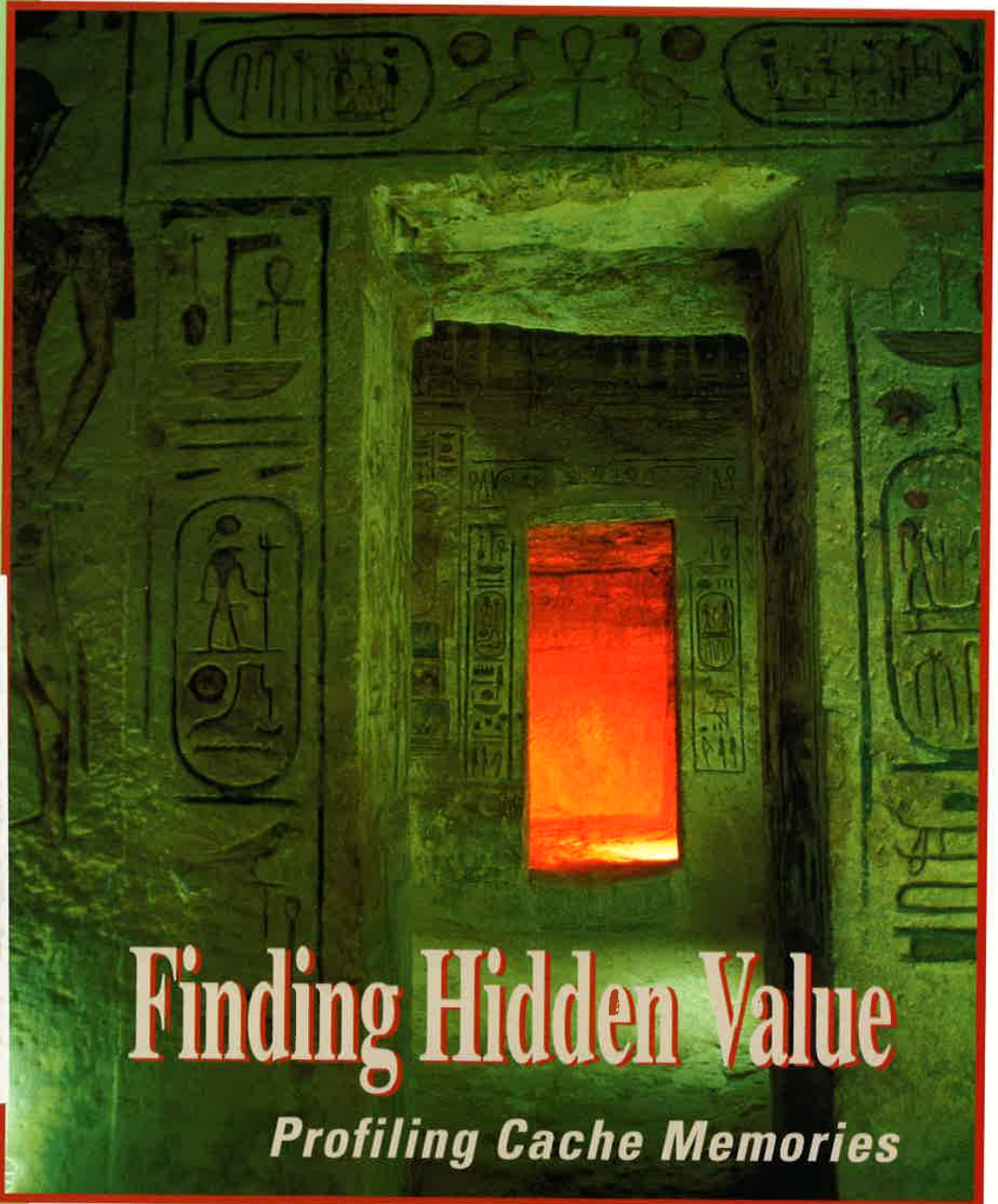


OCTOBER 1994

COMPUTER

Innovative technology for computer professionals



Finding Hidden Value

Profiling Cache Memories

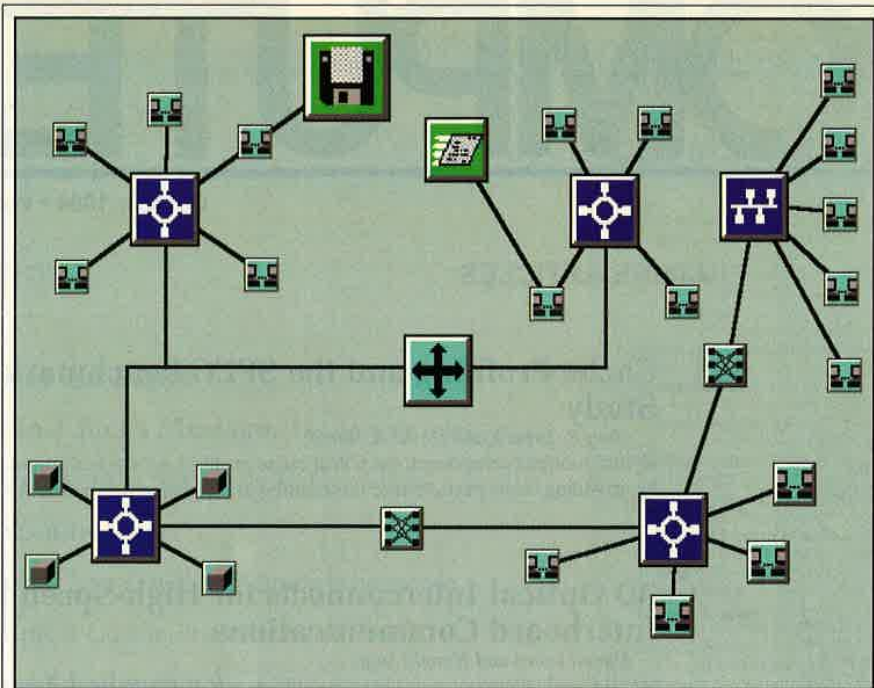
Also: 3D Optical Interconnects • Displaying 3D Images • Mosaic and WWW • Porting Ada • Software Control • New Department, Computer Law • **Special Book Offer, See Back Cover**

Join the
Computer 100.
See p. 6.

***** 3-DIGIT 205

 1617753 OCT94 300
 LIBRARY OF CONGRESS
 GROUP PERIODICALS REGISTRATION
 WASHINGTON DC 20540
 00521

IEEE COMPUTER SOCIETY
 THE INSTITUTE OF ELECTRICAL AND
 ELECTRONICS ENGINEERS, INC.



A typical application. **You** show a picture of **your** model with **your** icons — your graphical, animated results are easier to understand.

After you've tried new MODSIM II,
 other object-oriented languages will seem primitive
 Free trial and, if you act now, free training

MODSIM II is a modern high-level, object-oriented programming and simulation language with these important benefits:

- **Simulation.** All the features you need to do simulation and automatically collect statistics are built-in.
- **Object-oriented.** The powerful mechanisms that help you develop reusable, easily maintainable code.
- **Your program is portable.** Your programs are easier to read, and are portable.
- **Symbolic debugger.** Error detection simplified through steptracing, multiple breakpoints, and data viewing.
- **Graphics and animation.** You easily lay out your graphical user interface. You can draw icons and place them on a palette with no programming.

Free trial offer

The free trial contains everything you need to try MODSIM II on your computer. You receive everything you need for a complete evaluation.

For a limited time we also include free training.

Computers with MODSIM II

MODSIM II® runs on all popular workstations as well as PC's with Windows or OS/2. Your investment is preserved when you change computers, or operating systems.

MODSIM II is already in use at hundreds of installations throughout the world.

For immediate information

Call Hal Duncan at (619) 457-9681, Fax (619) 457-1184. In Europe, call Peter Holt, In The Netherlands, on 31 43 670780, Fax 31 43 670200. In the UK, call Nigel McNamara on 0276 671 671, Fax 0276 670 677.

Rush information on MODSIM II.

- Yes, I want to learn the reasons for the growing popularity of MODSIM II.

Special offer — return the coupon today and we will include a free course enrollment worth \$750.

Name _____
 Organization _____
 Address _____
 City _____
 Telephone _____
 Computer _____ O/S _____

- Send details on your University offer.

Return to:
 CACI Products Company
 3333 North Torrey Pines Court
 La Jolla, California 92037
 Call (619) 457-9681 Fax (619) 457-1184

In Europe:
 CACI Products Division
 MECC Business Center
 G. Martino Laan 85, 8th Floor
 6229GS Maastricht, The Netherlands
 Call 31 43 670780 Fax 31 43 670200

In the UK:
 CACI Products Division
 Coliseum Business Centre
 Watchmoor Park, Riverside Way
 Camberley, Surrey GU15 3YL, UK
 Call +44 (0)1276 671 671 Fax +44 (0)1276 670 677

CACI Products Company

MODSIM II is a registered trademark of CACI Products Company. ©1993

COMPUTER

October 1994 • Vol. 27, No. 10

**Membership
Magazine of the
IEEE
Computer
Society**

Circulation: *Computer* (ISSN 0018-9162) is published monthly by the IEEE Computer Society. IEEE Headquarters, 345 East 47th St., New York, NY 10017-2394; IEEE Computer Society Publications Office, 10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720-1264, phone, (714) 821-8380; IEEE Computer Society Headquarters, 1730 Massachusetts Ave. NW, Washington, DC 20036-1903. Annual subscription included in society member dues. Non-member subscription rates: available upon request. Single copy prices: members \$10.00; nonmembers \$20.00. This magazine is also available in microfiche form.

Postmaster: Send undelivered copies and address changes to *Computer*, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08855. Second class postage is paid at New York, New York, and at additional mailing offices. Canadian GST #125634188. Printed in USA.

Copyright and reprint permission: Copyright © 1994 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limits of US copyright law for private use of patrons: (1) those post-1977 articles that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Dr., Danvers, MA 01923; (2) pre-1978 articles without fee. For other copying, reprint, or republication permission, write to Permissions Editor, *Computer*, 10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720-1264.

Editorial: Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author's or firm's opinion. Inclusion in *Computer* does not necessarily constitute endorsement by the IEEE or the Computer Society. All submissions are subject to editing for style, clarity, and space.



FEATURE ARTICLES

15 Cache Profiling and the SPEC Benchmarks: A Case Study

Alvin R. Lebeck and David A. Wood

A vital tool-box component, the CProf cache profiling system lets programmers identify hot spots by providing cache performance information at the source-line and data-structure level.

27 3D Optical Interconnects for High-Speed Interchip and Interboard Communications

Ahmed Louri and Hongki Sung

Metal-based communications between subsystems and chips has become the limiting factor in high-speed computing. Maturing optics-based technologies offer advantages that may unplug this bottleneck.

38 Displaying 3D Images: Algorithms for Single-Image Random-Dot Stereograms

Harold W. Thimbleby, Stuart Inglis and Ian H. Witten

A new, simple, and symmetric algorithm can be implemented that results in higher levels of detail in solid objects than previously possible with autostereograms.

COMPUTING PRACTICES

49 Mosaic and the World-Wide Web

Ronald J. Vetter, Chris Spell and Charles Ward

This introduction to Mosaic and the World-Wide Web tells how to obtain a Web browser and to navigate on the Web.

58 Porting Ada: A Report From the Field

Joseph G. Skazinski

With the constant barrage of new systems, porting software applications is inevitable. This article takes a look at the problems involved in porting an Ada/C application.

COMPUTING MILIEUX

65 High-Pressure Steam Engines and Computer Software

Nancy G. Leveson

How can we get the benefits of computers without increasing risk? Studying parallels in the early development of high-pressure steam engines and of software engineering can help.

DEPARTMENTS

- 5 Letters to the Editor
- 6 Editor-in-Chief's Message: The next big thing
- 8 Industry Trends: NetWare's RISCy business
- 12 Article Summaries
- 74 Computer Law: Intellectual property protection
- 76 Hot Topics: Customizing information, Part 2
- 79 Software Challenges: Paperwork
- 81 New Products
- 86 Product Reviews: 3D CAD; Windows X servers
- 92 Standards: Information infrastructure committee
- 94 Update: Federal R&D opportunities
- 96 Call for Papers/Calendar
- 104 Career Opportunities
- 110 Media Reviews: NP-complete problems; Tcl/Tk
- 112 Open Channel: 12 parallel programming principles

Computer Society Information, 4; Membership Application, 48A; Advertiser/Product Index, 80; Reader Service Card, 80A; Change-of-Address Form, 64.

In the next issue: Associative processing and processors

On CS On-Line: See p. 93 for a list of items now available on the Computer Society's infoserfer.



Cover image: The tomb of Ramses II; Stock Imagery © 1993, G. Adams
Cover design: Chris Patterson

EDITOR-IN-CHIEF

Ted Lewis, Naval Postgraduate School
lewis@cs.nps.navy.mil, (408) 656-2449

ARTICLE SUBMISSIONS

Articles are limited to 6,000 words, with each figure/table counted as 250 words, and 12 references. Send enquiries and requests for author guidelines to computer@cs.nps.navy.mil. Submit 10 copies of the manuscript to Computer, Code CS, Spangale Hall, Rm. 513, Naval Postgraduate School, Monterey, CA 93943-5118. For special-issue proposals, send ASCII text to jf@cacs.usl.edu or send 20 printed copies to Joan Francioni at the address below.

AREA EDITORS

Theme Issues: Joan Francioni, Univ. of Southwestern Louisiana, CS Dept., Lafayette, LA 70504-1771
jf@cacs.usl.edu, (318) 231-6602

AI/Multimedia: Beverly Woolf, Univ. of Massachusetts
bev@cs.umass.edu, (413) 545-4265

Algorithms: Fan Chung, Bell Communications
frkc@bellcore.com, (201) 829-4324

Computer Integrated Manufacturing: Dieter Haban
Daimler Benz AG, haban%fuzi.uucp@germany.eu.net

Computing Milieux: Kathleen Swigger, Univ. of N. Texas
kathy@ponder.csci.utd.edu, (817) 565-2817

Computing Practices: Paul Oman, Univ. of Idaho
oman@cs.uidaho.edu, (208) 885-7219

Concurrent Systems: Tzila Elrad, Ill. Inst. of Technology,
cseftrad@minna.acc.ill.edu, (312) 567-5142

Database Systems: Ophir Frieder, George Mason Univ.
ophir@cs.gmu.edu, (703) 993-1530

Graphics and User Interfaces: Arie Kaufman
SUNY Stony Brook, ari@cs.sunysb.edu, (516) 632-8441

High-Performance Computing: Cherri Pancake
Oregon State Univ., pancake@cs.orst.edu, (503)737-2109

Networking: Louise Moser, Univ. of Calif., Santa Barbara
moser@ece.ucsb.edu, (805) 893-4897

Object Technology: Doris Carver, Louisiana State Univ.
carver@bit.csc.lsu.edu, (504) 388-3901

Parallel Computing: Vicki Allan, Utah State Univ.
allanv@cs.usu.edu, (801) 797-2022

Judith Schlesinger, Supercomputing Research Center
judith@super.org., (301) 805-7455

Real-Time Systems: Lui Sha, Carnegie Mellon Univ.
sha@sei.cmu.edu, (412) 268-3272

Software Metrics: Howard Rubin, Hunter College
71031.377@compuserve.com, (914) 764-4931

Systems Architecture: Laxmi Bhuyan, Texas A&M Univ.
bhuyan@cs.tamu.edu, (409) 845-9640, (409) 847-8578 fax

VLSI Design: Donald Bouldin, Univ. of Tennessee
bouldin@sun1.engr.utk.edu, (615) 974-5444

DEPARTMENT EDITORS

Computer Law: Irah Donner, Lowe, Price, LeBlanc, and Becker
74521.3320@compuserve.com, (703) 684-1111

Hot Topics: Ronald D. Williams, Univ. of Virginia, Charlottesville
rdw@virginia.edu, (804) 924-7960

Media Reviews: Janusz Zalewski, Univ. of Texas, Permian Basin
zalewski_j@gusher.pb.utexas.edu, (915) 367-2219

News: Nan C. Schaller, Rochester Inst. of Tech.
ncs@cs.rit.edu, (716) 475-2139

Open Channel: Will Tracz, Loral Federal Systems
tracz@lfs.loral.com, (607) 751-2169

Product Reviews: Jogikal Jagadeesh, Ohio State Univ.
jag@ohstphrm.pharmacy.ohio-state.edu, (614) 292-1903

Software Challenges: Capers Jones, Software Productivity
Research, capers@spr.com, (617) 273-0140

Standards: Norman Schneidewind, Naval Postgraduate School
schneidewind@nps.navy.mil, (408) 656-2719

STAFF

IEEE Computer Society Publications Office
10662 Los Vaqueros Circle, PO Box 3014
Los Alamitos, CA 90720-1264
t.seaborn or m.potes@computer.org, (714) 821-8380

Managing Editor: Marilyn Potes
Associate Editor: Christine Miller

Articles Editor: Scott Hamilton

Staff Editors: Bob Carlson, Chuck Governale,
Louise Burnham

Production: Chris Patterson, Ken Duckworth,
Jeff Messecar

Publisher: True Seaborn

Assistant Publisher: Matt Loeb

Ed. Dir., CS Magazines: Marilyn Potes

Memb./Circ. Promotion Mgr.: John Gill

Advertising Manager: Heidi Flex

Advertising Coordinator: Marian Tibayan

MAGAZINE OPERATIONS COMMITTEE

James J. Farrell III (chair), Valdis Berzins,
B. Chandrasekaran, Carl Chang, Dante Del Corso,
Ramesh Jain, John A.N. Lee, Ted Lewis, Michael J. Quinn,
Ahmed Samah, Ken Wagner, Peter R. Wilson

PUBLICATIONS BOARD

Barry W. Johnson (chair), Jon T. Butler, James J. Farrell III,
Tadao Ichikawa, Mary Jane Irwin, Raymond Miller,
Michael C. Mulder, Theo Pavlidis, John P. Fliganati,
Harold Stone, Ronald D. Williams

COMPUTING PRACTICES



Back in the late 1970s, I was predicting that a more user-friendly operating system would quickly supplant Unix. But I hadn't reckoned on MIT's X Windows. More recently, I suggested that the Internet was passé, soon to be overshadowed by advances in multimedia office environments. Oops, wrong again. Mosaic is to Internet what X Windows was to Unix.

Our first article, "Mosaic, HTML, and the World-Wide Web," by Ronald Vetter, Chris Spell, and Charles Ward, provides an introduction to the Mosaic browser, a GUI for this net-based hypertext system. If you're a net user not (yet) familiar with Mosaic, this article will get you started. If you're already a Mosaic user, you'll still appreciate this overview.

Our second article effectively embodies the original intent of this column. Months ago we ran an article describing an Ada tool for analyzing system dependencies. Joseph Skazinski wrote me to say that while he found the article interesting, it didn't fully describe the difficulties of porting Ada. I asked for examples, and he sent a short draft detailing his experiences. After some editorial give and take, I'm pleased to present his "Porting Ada: A Report From the Field." It's a nuts-and-bolts article about porting a large military application across several platforms. Even if you're not working in Ada, I think you'll find some of Joseph's suggestions useful.

Oh, and thanks to all of you who have written to me about the last two issues. Your feedback will ensure that Computing Practices stays in touch with the working engineer. As usual, feel free to contact me about what you liked and didn't like here in Practices.

— Paul Oman

Mosaic and the World-Wide Web

Ronald J. Vetter

North Dakota State University

Chris Spell and Charles Ward

University of North Carolina at Wilmington

The World-Wide Web, one of the newest information services on the Internet, uses hypertext links to other textual documents or files.¹ With this method, users can click on a highlighted word or words in the text to provide additional information about the selected word(s). Users can also access graphic pictures, images, audio clips, or even full-motion video through hypermedia, an extension of hypertext.

The World-Wide Web project was started by CERN (European Center for Nuclear Research) in an attempt to build a distributed hypermedia system.² Users access WWW through a "browser program" that can read and fetch documents locally as well as from sites around the world via the Internet. Browsers access files using FTP, NNTP, Gopher, and several other Internet-based protocols (see the glossary of terms). Browser clients can also search remote documents and databases if the associated server at that site has built-in search capabilities.

Browsers can be line-oriented (to work with traditional terminals) or graphics-oriented (to work with more modern graphics display terminals). One of the most popular graphics-oriented browsers is Mosaic, which was developed at the National Center for Supercomputing Applications (NCSA) as a way to graphically navigate the WWW. The first version of NCSA's Web browser, a WWW client for Unix with X Windows called XMosaic, was made available to the Internet community in 1993. Due primarily to its easy-to-use graphical user interface (GUI), XMosaic soon became the most popular interface to the Web.³

Mosaic is a system for wide-area distributed asynchronous collaboration and hypermedia-based information discovery and retrieval. Mosaic browsers are currently available for Unix workstations running X Windows, PCs running Microsoft Windows, and Macintosh computers. Mosaic can access data in WWW servers, Wide Area Information Servers (WAIS), Gopher servers, Archie servers, and several others. Figure 1 shows an example Mosaic screen (the home page) from the University of North Carolina at Wilmington.

Many commercial and government organizations and universities are migrating to Mosaic because of its many useful features and capabilities.^{2,3} It has a consistent mouse-driven graphical interface and can display electronic data in a variety of fonts and styles. Mosaic reads SGML (Standard Generalized Markup Language) and

Glossary of terms

Archie — A system that allows searching of indexes of files available on public servers by anonymous FTP on the Internet.

CERN — European Center for Nuclear Research.

FTP — File Transfer Protocol.

GIF — Graphics Interchange Format.

Gopher — A distributed information service that makes available hierarchical collections of information across the Internet. Both the server and client software are available as public domain software from the University of Minnesota.

HTML — HyperText Markup Language.

HTML+ — HTML extension that supports new features.

HTTP — HyperText Transfer Protocol, the native WWW protocol.

ISO — International Standards Organization.

Mosaic — A system for wide-area distributed asynchronous collaboration and hypermedia-based information discovery and retrieval. Mosaic browsers are high-level

graphical user interfaces (GUIs) currently available for Unix workstations running X Windows, PCs running Microsoft Windows, and the Apple Macintosh.

NCSA — National Center for Supercomputer Applications.

NNTP — Network News Transfer Protocol.

PPP — Point-to-Point Protocol, which allows a computer to use the TCP/IP protocols with a standard telephone line and a high-speed modem. PPP is replacing SLIP (see below) as a standard.

SGML — Standard Generalized Markup Language.

SLIP — Serial Line Internet Protocol, which allows a computer to use TCP/IP (Internet) protocols with a standard telephone line and a high-speed modem.

TCP/IP — Transmission Control Protocol/Internet Protocol.

Telnet — Lets users connect to any remote system on which the user has an account.

URL — Uniform Resource Locator.

HTML (HyperText Markup Language) text, enabling information to be presented in an attractive way. Mosaic also supports a wide variety of basic forms elements, such as fields, check boxes, and radio buttons, as well as characters from different languages (as defined by the ISO 8859 standard), including French, German, and Hawaiian. Finally, Mosaic supports graphics of up to 256 colors (in GIF or X bitmap format) as well as digital audio and video.

Users can extend Mosaic's functionality by creating custom servers and by letting other applications control its display remotely. Screen contents can be broadcast to a network of users running multiplatform groupware, such as NCSA's Collage, using NCSA's Data Transfer Mechanism (DTM). Mosaic users can traverse hypermedia links to provide basic network services (for example, FTP, Gopher, Telnet, NNTP, and WAIS) and store those links in "hotlists" for quick reference. And with the additional features available in HTML+, Mosaic users can "check out" Web documents for revision or for local or remote annotation.

HTML and HTML+

The Web uses HTML for creating and recognizing hypermedia documents.³ This is a simple markup system, related

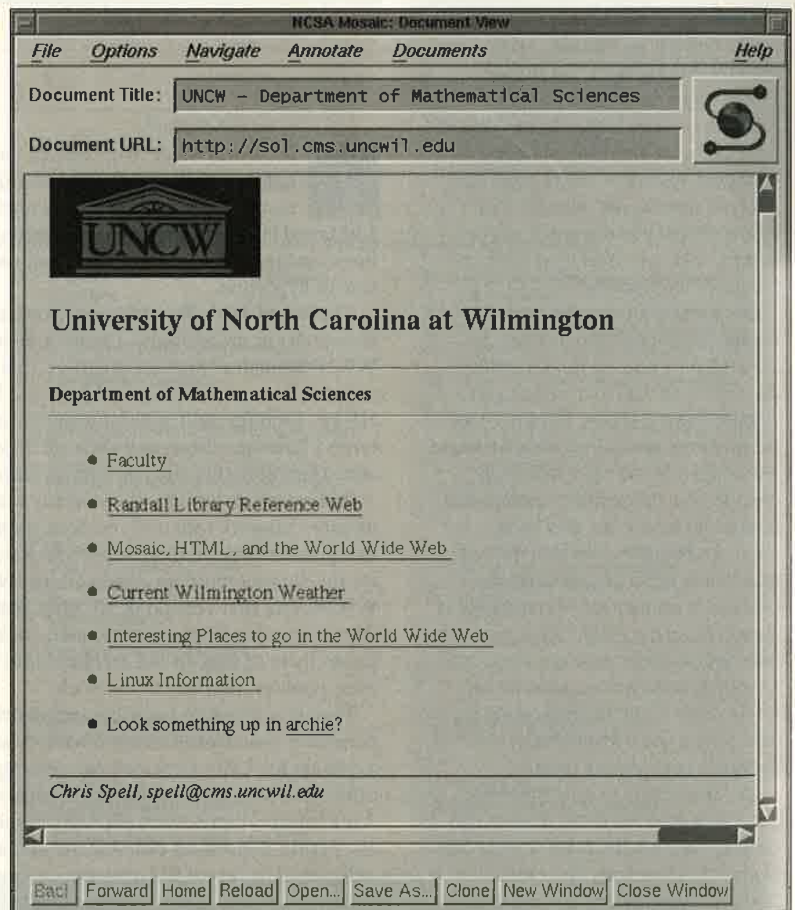


Figure 1. Sample Mosaic screen for the home page of the University of North Carolina at Wilmington.

to the Standard Generalized Markup Language (SGML), used to format hypertext documents. (On the Web, these documents usually carry the ".html" suffix.) HTML code describes how textual elements (like paragraphs, lists, numbered and bulleted lists, descriptive lists, and quoted paragraphs) will be displayed. The current HTML standard supports basic hypermedia document creation and layout, but this capability is still very limited. The HTML+ extension supports interactive forms, defined "hotspots" (hyperlinks) in images, formatted tables, and more versatile document layout, formatting options, and styles.

Writing HTML documents. Most HTML documents are written using a standard text editor — for example, *vi* in a Unix environment — in a plain text format. Although HTML code is a tedious language to write, some authoring tools can now convert plain text to HTML automatically.

One way to learn how to write HTML documents is to look at how others have created their documents. You can do this by using the "source" button (available on most Mosaic browsers) to view the HTML code for a document or page you find particularly interesting. Most HTML statements are fairly straightforward. For example, Figure 2 shows the HTML code for the Mosaic screen image depicted in Figure 1. The code corresponds closely to the final graphic document. If you intend to write HTML code, we recommend that you read the on-line Mosaic document "A beginner's guide to HTML," which is available on the Web (see sidebar on HTML editors and converters).

HTML editors. HTML documents can be created with WYSIWYG (what you see is what you get) editors or with editors that simply assist you by allowing you to select the desired markup tags from a menu. HTML editors are easier to use and interactively display the resulting HTML document on-screen. For example, current versions of the Emacs editor have an "HTML mode" to assist users with writing statements. There is also an editor for Microsoft Windows called HTML Assistant. For X users, tkWWW

```

<TITLE>UNCW - Department of Mathematical Sciences</TITLE>

<H1>
University of North Carolina at Wilmington
</H1>

<hr> <b>
Department of Mathematical Sciences</b>
<hr> <ul>

<p> <li>
<a href="faculty.html">
Faculty
</a>

<p> <li>
<a href="/randall/ref.html">
Randall Library Reference Web
</a>

<p> <li>
<a href="mosaicpaper.html">
Mosaic, HTML, and the World Wide Web
</a>

<p> <li>
<a href="newweather.html">
Current Wilmington Weather
</a>

<p> <li>
<a href="OtherLinks.html">
Interesting Places to go in the World Wide Web
</a>

<p> <li>
<a href="/linux/linux.html"> Linux Information
</a>

<p> <li>
Look something up in <A HREF="/cgi-bin/archie">archie</A>? <P>
<p> <li> <p> </ul> <p>

<address>
Chris Spell, spell@cms.uncwil.edu
</address>

```

Figure 2. HTML code for the Mosaic screen shown in Figure 1.

supports WYSIWYG HTML editing; and since it is a browser, you can try out links immediately after creating them. HotMetal is a professional HTML Plus editor for X and Microsoft Windows. For Macintosh users, the BBEdit HTML extensions allow the BBEdit and BBEdit Lite text editors to conveniently edit HTML documents.

Navigating the Web

Information currently accessible through the Web includes information served through Telnet, Gopher, WAIS, FTP, Usenet news, Archie, and anything in the form of Unix man pages or hypermedia documents. Mosaic lets users effectively navigate the Web and access the



wide variety of information databases associated with these existing Internet resources.

Uniform Resource Locators. A Uniform Resource Locator (URL) refers to the format used by WWW documents to locate other files and is currently a draft standard for specifying an object or resource on the Internet. A URL gives the type of resource being accessed (for example, Gopher or WAIS) and the path of the file. The format used is: *scheme://host.domain[:port]/path/filename*.

The first part of the URL, the scheme, specifies the access method. Several keywords can be specified for the scheme:

- FTP — retrieves a file on your local system or on an anonymous FTP server,
- HTTP — retrieves a file on a WWW server (the native WWW protocol),
- Gopher — retrieves a file on a Gopher server,
- WAIS — retrieves a file on a WAIS server,
- News — reads the latest Usenet news, or
- Telnet — uses the telnet protocol to connect to some site.

The part of the URL after the colon is interpreted with regard to the specific access method. In general, two slashes after the colon indicate a machine name or, optionally, a port. (For additional information, see "A Beginner's Guide to URLs," <http://www.ncsa.uiuc.edu/demoweb/url-primer.html>.) The last part of the URL describes the location-specific path and filename of the desired document. Some example URLs are:

- <ftp://ftp.uu.net/info/README>
- <ftp://ftp.uu.net/usenet>
- <http://info.cern.ch:80/default.html>
- <news:alt.hypertext>
- <telnet://dra.com>

HTML editors and converters

BBEdit HTML — A Macintosh editor for HTML documents obtained by accessing <http://www.uji.es/bbedit-html-extensions.html>. You can also obtain the extensions package by anonymous FTP from <sumex-air.stanford.edu> in file [info-mac/text/bbedit-html-ext-b4.hqx](ftp://sumex-air.stanford.edu/info-mac/text/bbedit-html-ext-b4.hqx).

Emacs editor — Current version Emacs have an HTML mode to assist users with writing HTML statements (<ftp://ftp.ncsa.uiuc.edu/Web/Elisp/html-mode.el>).

HTML — An introductory reference, "A Beginner's Guide to HTML," is available at <http://www.ncsa.uiuc.edu/General/Internet/WWW/HTMLPrimer.html>. The HTML primer is available at <http://www.vuw.ac.nz/who/Nathan.Torkington/ideas/www-html.html>.

HTML Assistant — A Microsoft Windows editor with features to assist in the creation of HTML documents can be obtained by anonymous FTP from <ftp.cs.dal.ca/htmlasst>.

HTML+ — To learn more about HTML+, you can examine the ASCII text of a draft specification for it at <ftp://ds.internic.net/internet-drafts/draft-raggett-www-html-00.txt>; or a PostScript version of the same at <ftp://ds.internic.net/internet-drafts/draft-raggett-www-html-00.ps>

LaTeX2HTML — Converts LaTeX documents to HTML (<http://cbl.leeds.ac.uk/nikos/latex2html/doc/latex2html.html>).

Mail2HTML — Converts electronic mailboxes to HTML documents (<ftp://info.cern.ch/pub/www/dev>).

NCSA's list of filters and editors — This document mentions several editors, including two for Microsoft Windows (<http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/faq-software.html#editors>).

RTF2HTML — Converts Rich Text Format (RTF) to HTML (<file://oac.hsc.uth.tmc.edu/public/unix/WWW>).

HyperText Transfer Protocol. HTTP is the protocol employed between the server and client.⁴ It requires only a reliable connection-oriented transport service, typically TCP/IP. The client establishes a connection with the server and sends a request containing the word GET, a space, and the partial URL of the node to be retrieved, terminated by the carriage return and line feed characters. The server responds with the node contents, which consist of HTML text documents. The end of the contents is signaled by the server closing the connection.

Ways to spin a Web. Creating a Web can be as simple as pointing a hyperlink from an existing Web document to your existing data. Making it public involves running the FTP or HTTP daemon. Any file available by anonymous FTP can be immediately linked into the Web. This method requires only a small start-up effort and allows contributions to be made network accessible via the Internet in a very short period of time.

You can also create a Web by writing your own HTML code and pointing it to databases of interest. Data may come from local archives or be stored elsewhere on the Internet. This requires familiarity with the HTML language and the information resources currently available on the Internet. Existing files (for example, a LaTeX document) can be converted to HTML automatically by using a set of special software tools.⁵

When creating Web documents for public access, we have found the following guidelines to be useful.

- Remember that most users accessing information over the Internet are browsing and are therefore unwilling to wait long for a document to appear on their screen. You should use graphics only to catch the reader's attention. Keep the size of images on the home page small; they should be used as simple icons that show what the object points to. Wait until the user selects the icon to incorporate high-quality (and large) images into your documents. This will let users access the home page quickly and will ease server and network load. Although each image takes time to process and slows the display, using a particular image multiple times in a

document causes very little performance degradation compared to using it once.

- Documents transmitted over a network will not display as quickly as documents will locally, where you typically develop them. Before making any document public, be sure to test the document's remote access time to ensure that it is reasonable.

- Always provide an estimate of the number of bytes that will be sent when a user selects a given "hotword" (especially for high-resolution images and digital audio and video). This helps users decide whether to select a given hotword.

- Do not mix a large number of different fonts, colors, and so forth on a single display screen. Use good graphic arts design principles. This is especially important if the screens will be used for educational purposes and projected onto an overhead screen.⁶ For example, certain colors that look good together on screen and are easily distinguishable may not be so distinguishable when sent through a projection panel and displayed with an overhead projector. We have provided some guidelines to help developers of multimedia presentations avoid some of the common problems associated with the projection of these materials.⁶

- Do not say "Click Here." It is redundant and only clutters the screen. Hotwords are normally a different color or font, so it is obvious (especially to fre-

quent users) that "clicking here" will cause a hyperlink to be traversed.

- Put a reference (for example, an electronic mail address) at the bottom of every page telling whom to contact about the document. This can be accomplished by using the HTML Address Tag construct.

- Do not make your pages too long. Rather, create a hierarchical set of menus that group items into meaningful categories. That way only those items of interest will have to be selected.

How does the WWW compare to Gopher and WAIS?

While all three of these information presentation systems are client-server based, they differ in terms of their data models. In Gopher, data is either a menu, a document, an index, or a Telnet connection. In WAIS, everything is an index, and everything that is returned from the index is a document. In WWW, everything is a hypertext document, which may be searchable if it contains text. This means that WWW can represent Gopher and WAIS data models as well as provide extra functionality. Only WWW has hyperlinks.

Although text is supported by all three systems, WWW provides facilities for dis-

playing "richer" text (for example, headings, lists, and emphasized text) in a standardized way. Image, audio, and video data are all supported by external viewers (for example, MPEGplay). There is little direct support for application-specific data in any of the three systems, and all use different methods for expressing type and encoding. Also, all three systems have little support for controlling the presentation of nontext data. For example, backdrops and synchronization in time are not supported, and buttons are only supported in a limited way. (See Adie² for a more comprehensive evaluation.)

Limitations. Perhaps the greatest limitation of current distributed hypermedia systems is that there is no simple way to find out what information has changed, what information is new, or even what information is out there on the Internet. Mosaic users commonly complain about difficulties finding resources on a particular topic or subject. Unless the user has extensive knowledge of the Internet, it is hard to locate resources of interest on the WWW. A related problem is remembering where resources were found when trying to access them again (hotlists are useful, but it is better to annotate the location and description of the resource on a notepad or set of index cards so that it can later be recalled). Another problem is that there are simply too many links,

Getting started: How to access the Web¹

To access the Web, you first need an account on a machine that is connected to the Internet. All WWW features can be accessed with client/browser software from this machine. To obtain a browser for your machine, refer to the discussion below. Many of the browsers discussed here require that you have SLIP, PPP, or other TCP/IP networking to your computer. SLIP or PPP can be accomplished over phone lines, but only with the active cooperation of your network provider or educational institution. If you only have normal dial-up access, your best option is to run a browser on the system you call directly or from a system that can be connected to.

Browsers accessible via Telnet. Several browsers are available over the Internet through Telnet. An up-to-date list of

these is available on the Web as <http://info.cern.ch/hypertext/WWW/FAQ/Bootstrap.html>. For example,

- **Info.cern.ch** — No password is required. This is in Switzerland, so continental US users might be better off using a closer browser.
- **ukanaix.cc.ukans.edu** — A full-screen browser "Lynx" that requires VT100 terminal emulation. Log in as **www**.
- **www.njit.edu** — A full-screen browser in New Jersey Institute of Technology. Log in as **www**.
- **vms.huji.ac.il** — A dual-language Hebrew/English database at Hebrew University of Jerusalem in Israel, with links to the rest of the world. Log in as **www**.
- **sun.uakom.cs** — A machine in Slovakia. Log in as **www**.
- **fserv.kfki.hu** — A machine in Hungary. Log in as **www**.

Continued on overleaf

Table A. Mosaic browsers.

Operating system	Program	FTP Availability	Comment
Windows	Cello	ftp.law.cornell.edu:/pub/LII/cello	
	Mosaic for Windows	ftp.ncsa.uiuc.edu:/PC/Mosaic	
Macintosh	Mosaic for Macintosh Samba	ftp.ncsa.uiuc.edu:/Mac/Mosaic info.cern.ch:/pub/www/bin/mac	
Amiga	AMosaic	max.physics.sunysb.edu:/pub/ amosaic	
NextStep	OmniWeb	ftp.omnigroup.com:/pub/software; for information, http://www. omnigroup.com/	The browsers listed here are native NextStep applications. Next systems can also run X-based browsers using one of the widely used X-server products for the Next.
	CERN's Browser-Editor	info.cern.ch:/pub/www/src	Allows WYSIWYG hypertext editing. Requires NextStep 3.0.
X/Dec- Windows	Mosaic for X	ftp.ncsa.uiuc.edu:/Web	Unix browser using X11/ Motif; compiled for several popular platforms.
	Mosaic for VMS	ftp.ncsa.uiuc.edu:/Mosaic	NCSA Browser using X11/DecWindows/Motif.
	tkWWW Browser/	harbor.ecn.purdue.edu:/pub/tcl/ extensions/tkwww (followed by an extension dependent on the current version)	Editor for X11 supports WYSIWYG HTML editing.
	Viola for X (Beta)	ftp.ora.com:/pub/www/viola	Viola has two versions for Unix/X: one using Motif, one using Xlib (no Motif). Handles HTML+ forms and tables. Has extensions for multiple columning and collapsible/expandable lists. Client-side document included.
	Chimera	ftp.cs.unlv.edu:/pub/chimera	Unix/X Browser using Athena (doesn't require Motif).
Text-mode Unix and VMS browsers	Line Mode Browser	info.cern.ch:/pub/www/src	This program gives WWW access to anyone with a dumb terminal. A general-purpose information retrieval tool.
	Lynx full-screen	ftp2.cc.ukans.edu:/pub/WWW/lynx	This hypertext browser for VT100 terminals uses full browser screen, arrow keys, highlighting, etc.
	perlWWW	archive.cis.ohio-state.edu:/pub/ w3browser	A tty-based browser written in Perl.
	VMS client	vms.huji.ac.il:/www/vms_client	A full-screen client based on VMS's SMG screen management routines.

Obtaining your own browser. The preferred method of accessing the Web is to run a browser yourself. Browsers are available for many platforms, both in source and executable forms. Table A lists some available browsers (a more extensive listing can be found at <http://info.cern.ch/hypertext/WWW/Clients.html>).

Setting up your own server. Browsers can obtain data from programs supplied by information providers. These programs can either be WWW servers that understand the HyperText Transfer Protocol, "gateway" programs that convert an existing information format to hypertext, or a non-HTTP server that WWW browsers can access — anonymous FTP or Gopher, for example. To learn more about WWW servers, you can consult a WWW server primer available at <http://www.vuw.ac.nz/who/Nathan.Torkington/ideas/www-servers.html>. Table B lists some servers currently available for Unix, Macintosh, Microsoft Windows, and VMS systems. (For more information on writing servers and gateways in general, see <http://info.cern.ch/hypertext/WWW/Daemon/Overview.html>.)

Creating a public home page. There are several things you can do to publicize your new HTML server or other offering:

- Submit it to the NCSA "What's New Page" at <http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/whatsnew.html>.
- Post it to the newsgroup `comp.infosystems.announce`. Please read the newsgroup first to get a feel for the contents. You can also crosspost to `comp.infosystems.www`.
- Submit it to the maintainers of various catalogs, such as the WWW Virtual Library <http://info.cern.ch/hypertext/DataSources/bySubject/Overview.html>.

Reference

1. "World-Wide Web Frequently Asked Questions," <http://siva.cshl.org/~boutell/www-faq.html>, June 1994.

Table B. Mosaic servers.

Operating System	Program	FTP Availability	Comment
Unix	NCSA httpd	ftp.ncsa.uiuc.edu:/Web/ncsa_httpd	
	CERN httpd	info.cern.ch:/pub/www/src	
	GN Gopher/HTTP server	URL http://hopf.math.nwu.edu	Serves WWW and Gopher clients. A good server for migrating from Gopher to WWW, although lacking the NCSA and CERN servers' script capabilities.
	Plexus	http://bsd.com/server/doc/plexus.html	Written in the Perl scripting language. Documentation available at the URL.
Macintosh	MacHTTP	http://www.uth.tmc.edu/mac_info/machttp_info.html	
MS Windows and Windows NT servers	HTTPS (Windows NT)	ftp://emwac.ed.ac.uk/pub/https	A server for Windows NT systems, both Intel- and Alpha-based.
	NCSA httpd (Windows)	ftp.ncsa.uiuc.edu:/Web/ncsa_httpd/contrib/whttp11a6.zip	The NCSA server has most features of the Unix version, including scripts.
	SerWeb	FTP://winftp.cica.indiana.edu/pub/pc/win3/winsoc/serweb03.zip	A simple, effective server for Windows.
VMS Servers	CERN HTTP	http://delonline.cern.ch/disk\$user/duns/doc/vms/distribution.html	A port of the CERN server to VMS.
	Region 6 Threaded HTTP Server	http://kcgl1.eng.ohio-state.edu/www/doc/serverinfo.html	A native VMS server that uses DEC-threads.



as Hall has pointed out in a recent article.⁷ Users of distributed hypermedia systems are often overwhelmed by the large number of possible links and become disoriented when moving between different application packages and information servers.

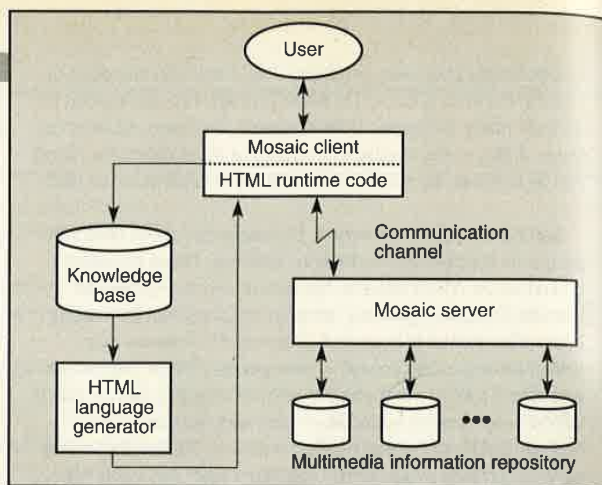
Steps are being taken to address some of these problems in Mosaic. Several resources now provide information on new and established Mosaic servers by topic. For example,

- The WWW Virtual Library⁸ is a good place to find resources on a particular subject, and
- What's New With NCSA Mosaic⁹ carries announcements of new servers on the Web.

More recently, several software tools have been developed that create automatic indexes you can search. Two such tools are WebCrawler (<http://www.bio-tech.washington.edu/WebQuery.html>) and WWW Worm (<http://www.cs.colorado.edu/home/mcbryan/WWW.html>). WebCrawler indexes the contents of the documents, whereas WWW Worm builds its index based on page titles and URL contents. Although tools that automatically traverse the Web are useful, tools that recursively search for information may not notice self-referencing links and may begin to return an infinite number of indexes.

Popularity. The WWW is gaining popularity as evidenced by the amount of network traffic (in bytes) across the National Science Foundation's North American Network (NSFNet).¹⁰ In fact, WWW traffic recently surpassed Gopher traffic, according to statistics produced by monitoring traffic on the Internet backbone. (Since WWW browsers can also access Gopher servers, the Gopher traffic may be somewhat inflated.) From January to August 1993, the amount of traffic attributed to Web use multiplied by 414 times, and usage of the server at CERN doubled every four months — twice the rate of Internet expansion.³ The Web is now ranked eighth of all network services in terms of sheer byte traffic as of May 1994 (Gopher is ranked tenth).

Figure 3. Mosaic extension model. An intelligent agent monitors user preferences to create a knowledge base that dynamically filters the hyperlinks presented to the user.



Extending Mosaic to support customized views

We are building an interface agent for Mosaic that helps users find the various types of Internet data available on the WWW. Our approach is similar to that undertaken by others who have developed autonomous agents to provide personalized assistance with meeting scheduling, electronic mail handling, electronic news filtering, and selection of entertainment.¹¹ The idea is to augment Mosaic with an intelligent interface agent that can learn the particular interests, habits, and preferences of individual users and then help them obtain the desired information in a way that reduces the number of links offered to the user. We accomplish this by supplying the interface agent with domain-specific background knowledge about user preferences and interests. At runtime, the interface agent uses this knowledge to find and recognize information of interest to the user. The agent must also be capable of acquiring knowledge on its own and building appropriate rules and facts in the knowledge base by monitoring user behavior over time (a form of machine learning).

Our approach aims at having Mosaic customize the link-following processes to suit a user's unique interests, habits, and access patterns. The system customizes the hypermedia environment supported by Mosaic by creating links dynamically, using rule-based algorithms or neural networks. This approach is much more flexible than many of today's closed hypermedia systems and still provides the user with the power to ask for more information when

needed rather than statically displaying all available links on a particular topic every time.

Figure 3 depicts the system model we propose. Initially, users (and agents) store information about particular interests in the knowledge base. The rules and facts are used as input to the HTML Language Generator. User-specific HTML code (and links) are then generated at runtime. This requires communication with other Mosaic agents and servers (much like Archie servers are used) to find the relevant information on a particular topic or subject.

The goal of our work is to demonstrate that an intelligent interface agent can be used to satisfactorily limit the number of links offered to the user while simultaneously providing users with a more effective interface to the WWW.

The World-Wide Web is still evolving at a rapid pace. Clearly, distributed hypermedia systems on the Internet will continue to be an active area of development in the future. We believe that the flexibility of the WWW design, its use of hyperlinks, and the integration of existing WAIS and Gopher information resources, make the WWW ideal for future research and study. We also realize that highly interactive multimedia applications will require more sophisticated tools than currently exist. The most significant issue that needs to be resolved is the mismatch between WWW system capabilities and user requirements in the areas of presentation and quality of service. The lack of adequate directory services will continue to plague WWW users, and future work on distributed hypermedia systems will need to address these issues.

References

1. E. Krol, *The Whole Internet: User's Guide and Catalog*, O'Reilly & Associates, Sebastopol, Calif., 1992.
2. C. Adie, "Network Access to Multimedia Information," Edinburgh University Computing Services, Edinburgh, Great Britain, Aug. 1993. Available at <ftp://ftp.edinburgh.ac.uk/pub/mmaccess>.
3. K. Hughes, "Entering the World-Wide Web: A Guide to Cyberspace," <ftp://taurus.cs.nps.navy.mil/pub/mby/world-wide-web-guide.ps.Z>, 1994.
4. T. Berners-Lee, "The HTTP Protocol as Implemented in W3," <ftp://info.cern.ch/pub/www/doc/http-spec.txt.Z>, Jan. 1992.
5. R. Brandwein and M. Sendall, "HTML Converters," <http://info.cern.ch/hypertext/WWW/Tools/Filters.html>, 1994.
6. R. Vetter, C. Ward, and S. Shapiro, "Experiences With Projected Multimedia Presentations," *Proc. Third Annual Conf. Multimedia in Education and Industry*, Assoc. for Applied Interactive Multimedia, Columbia, S.C., 1994.
7. W. Hall, "Ending the Tyranny of the Button," *IEEE Multimedia*, Vol. 1, No. 1, Spring 1994.
8. "The WWW Virtual Library," <http://info.cern.ch/hypertext/Data-Sources/bySubject/Overview.html>, June 1994.
9. "What's New with NCSA Mosaic," http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/whats_new.html, June 1994.
10. "World-Wide Web Growth," <ftp://nic.merit.edu>, 1994.
11. "Intelligent Agents," special issue, *Comm. ACM*, Vol. 37, No. 7, July 1994.

Ronald J. Vetter is an assistant professor in the Department of Computer Science and Operations Research at North Dakota State University in Fargo. His research interests include high-performance computing and communications, multimedia systems, and high-speed optical networks.

Vetter received the BS and MS degrees in computer science from North Dakota State University in 1984 and 1986, respectively, and his PhD in computer science from the University of Minnesota, Minneapolis, in 1992. Vetter is a member of the ACM and the IEEE Computer and Communication Societies.

Christopher G. Spell is the system administrator in the Mathematical Sciences Department at the University of North Carolina at Wilmington. His research interests include operating systems, communications, and multimedia systems. Spell received his BS in computer science from UNCW in 1993.

Charles R. Ward is assistant vice chancellor for academic affairs and professor of chemistry at the University of North Carolina at Wilmington. His research interests include the application of multimedia technology in chemical education.

Ward received the BS degree in chemistry from Manchester College in 1970, and the MS degree in chemistry from Purdue University in 1975. He received the PhD degree in science education from Purdue University in 1978. He is a member of the American Chemical Society and the Association for Applied Interactive Media.

Readers can contact the authors through Ronald Vetter, North Dakota State University, IACC Bldg., rm. 258, Fargo, N.D., 58105; e-mail rvetter@plains.nodak.edu.

October 1994



Call for Papers WET ICE 95

Fourth IEEE Workshop on Enabling Technologies:
Infrastructure for Collaborative Enterprises
April 20-22, 1995
Berkeley Springs, West Virginia

The Concurrent Engineering Research Center (CERC) at West Virginia University is seeking submissions for WET ICE '95, which is sponsored by the IEEE Computer Society. Papers reporting survey, original research, design and development, and applications of enabling technologies for collaboration are sought in the following areas:

virtual team support environments
mediators to support collaborative activities
information sharing in distributed systems
enterprise modeling
process modelling and characterization
integration of heterogeneous and legacy databases
applications of collaborative technologies (healthcare, software engineering, electronic commerce, etc.)
projects and team coordination
requirements, constraints
mobile computing
workflow tracking and management tools
networked collocation
tools for multi-media conferencing
capture of design intent and intelligent retrieval of corporate knowledge
enterprise integration frameworks

Papers should be no more than 25 typewritten, double-spaced, single-sided pages including all text, figures, and references. Papers should not have been published or be under submission currently for publication elsewhere. Manuscripts should have a title page that includes the title of the paper and the full name, affiliation, postal address, electronic address, telephone number, and fax number of each author. Authors also are encouraged to write a 300-word abstract and a list of keywords that identify the central issues of the paper. Papers can be submitted electronically as PostScript files (preferred mode) or by paper mail. Accepted papers will be included in the post-proceedings, to be published by IEEE Computer Society Press.

Five copies of the full manuscript due December 16, 1994
Notification of decisions February 15, 1995
Final version of the paper due May 15, 1995

General Chair: K. Joseph Cleetus, CERC, West Virginia University
Program Chair: John Callahan, CERC, West Virginia University

Program and Review Committee:

Chandrajit Bajaj, Purdue U.	Alex Schill, Technical U. of Dresden, Germany
Earl Craighill, SRI International	Dennis Sng, Nanyang Technological U., Singapore
Prasun Dewan, U. of North Carolina - Chapel Hill	Ram Duvvuru Sriram, NIST
Mark Fox, U. of Toronto	P. David Stotts, U. of North Carolina - Chapel Hill
Sumit Ghosh, Brown U.	George Trapp, West Virginia U.
Ted Goranson, SAIC	Robert Winner, Ctr. for High Performance Computing
Michael Huhns, MCC	
Raman Kannan, Monmouth College	
Vinay Kumar, Enterprise Integration Technologies	
Jintae Lee, U. of Hawaii	
John Lewis, GE Corporate R&D	
Felix Londono, U. EAFIT, Colombia	
James Purtilo, U. of Maryland - College Park	

Point of Contact:
John Callahan
CERC, P.O. Box 6506
West Virginia University
Morgantown, WV 26506
Phone: 304-293-7226 x286
Fax: 304-293-7541
Email: et-wkshp@cerc.wvu.edu