CIMEL: Constructive, collaborative Inquiry-based Multimedia E-Learning*

Glenn David Blank, Associate Professor, 610-758-4867, glenn.blank@lehigh.edu
William M Pottenger, Assistant Professor, 610-758-3454, pottenger@eecs.lehigh.edu
G. Drew Kessler, Assistant Professor, 610-758-4818, kessler@eecs.lehigh.edu
David Gevry, Research Assistant, 610-758-2727, drg2@lehigh.edu
Martin Herr, Undergraduate Developer, 610-758-6305, mah2@lehigh.edu
Harriet Jaffe, Research Assistant, 610-758-6305, haj4@lehigh.edu
Soma Roy, Research Assistant, 610-758-3737, sor4@lehigh.edu
Qiang Wang, Research Assistant, 610-758-4070, qiw3@lehigh.edu

Project web site, http://www.eecs.lehigh.edu/~cimel
*This project is funded by National Science Foundation (Grant No : EIA-0087977)
We welcome users and evaluators from other sites. If interested, please contact the first author.

We are developing a multimedia framework for constructive and collaborative, inquiry-based learning for Computer Science curricula, from introductory through graduate level courses. *Constructive* learning goes beyond learning by receiving knowledge, to learning by building systems, with immediate, visual feedback. *Collaborative* learning encourages students to interact with instructors and librarians, via both live links and remote-controlled "show me" sessions or by reviewing a multimedia FAQ of recorded "show me" sessions. *Inquiry-based* learning guides the student into pursuing exploratory research in a community of students and scholars. A reference librarian persona will suggest research topics, then help extract content from both traditional library resources as well as dynamically mined material, answer typical questions and help construct annotated bibliographies, reviews and research proposals. Our multi-track framework will allow students and instructors to select content according to their background, interests and learning goals. Within this general framework, we are developing content for two courses in the Computer Science curriculum, Introduction to Computing (CS1) at the freshman level (for both majors and non-majors) and Object-Oriented Software Engineering (OOSE) at the graduate level. In the fall of 2001, we gathered baseline evaluation data for the CS1 and OOSE courses, which our evaluators analyzed.

Since the project began in October 2000, we have developed a prototype for OOSE with Abstract Data Types as our sample content. Figure 1 provides a snapshot of the prototype, showing the content menu on the left, presentation area on the right and button bar on the lower right. The prototype is available at http://www.cse.lehigh.edu/~cimel/prototype.html. After introducing ADTs, the prototype includes an exercise to find emerging trends in topics related to students' research interests by mining relevant textual databases. The reference librarian persona guides the user through the process of detecting such trends in a "show me" session.

Our research in the area of textural data mining, we have focused on the automatic identification of emerging trends in textual data. Below, we discuss the integration of trend detection in the development of constructive, inquiry-based multimedia courseware. This integration requires sophisticated lesson tracking and context construction mechanisms.

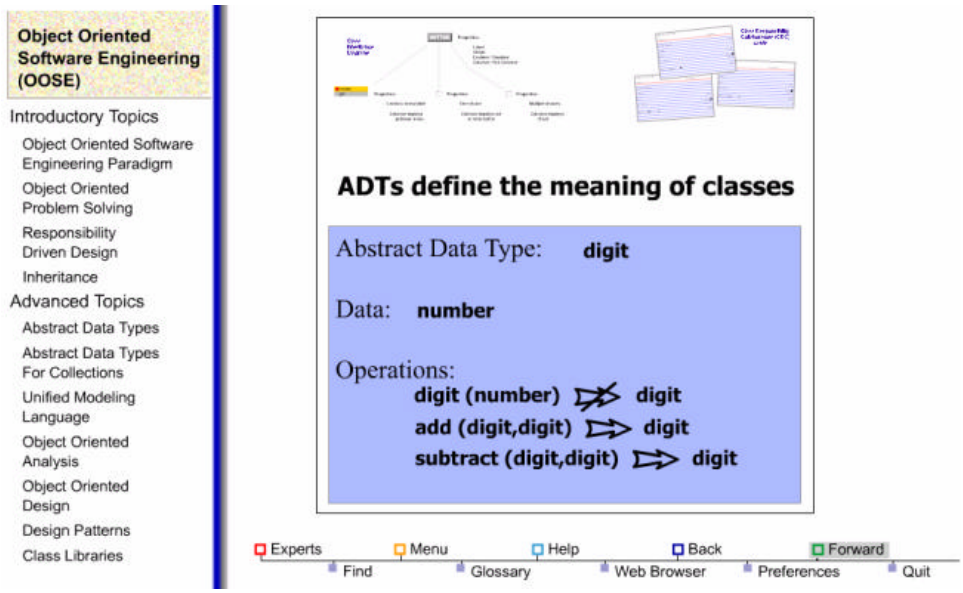

Figure 1: A snapshot of the CIMEL prototype

The research that is being conducted in this area is two fold. The first focus is to automatically track users as they move through multimedia lesson content and to gather statistics that will help determine how individual users as well as users as a group approach the lessons. Our goal is to use individual users' time-sensitive contexts to enhance their performance when conducting constructive, inquiry-based learning exercises that employ the Hot Topics Textual Data Mining System to uncover trends in a given field of study.

The motivation for this tracking research comes from our current work with user profiling based on temporal aspects of web access: how often a user visits a page and how long they stay on that page. The goal of the research is to link a user's temporal footprint with the semantic data of the documents that they view. This temporal footprint will allow us to automatically filter a model of the user's interests based on their history of access to the material. We will compare the performance of the Hot Topics Textual Data Mining System both with and without filtered models to determine the impact of employing a temporal context in constructive, inquiry-based learning exercises that incorporate trend detection in OOSE.

Since starting the project in October of 2000, we have accomplished the following:

- Completed development of a repository of over 2000 abstracts from the field of Object Oriented Software Engineering (OOSE). Based on domain-expert analysis of the documents in this repository, we have also developed a training set of approximately 250 instances of emerging and non-emerging topics for use in building models that will be deployed as part of the Hot Topics Textual Data Mining System.
- Completed development of an alpha version of the Hot Topics Textual Data Mining System (including a user interface for visualization of trends, as shown in Figure 2, a sample screenshot from the prototype).
- Completed design and rapid prototype of lesson tracking mechanisms that capture temporal footprint of user access to multimedia content developed using Flash.
- Begun design of tools and methods for evaluating the Hot Topics Textual Data Mining System as an integrated part of the multimedia courseware.

The figure below depicts the emergence of the topic "Components", a recent technique in OOSE. "Components" is contained in a cluster of related topics that grew in size starting in 1995 and continuing through 1998 (the year at which the mouse is pointing in the screenshot below). This is an indication that "Components" represents an emerging trend in OOSE.

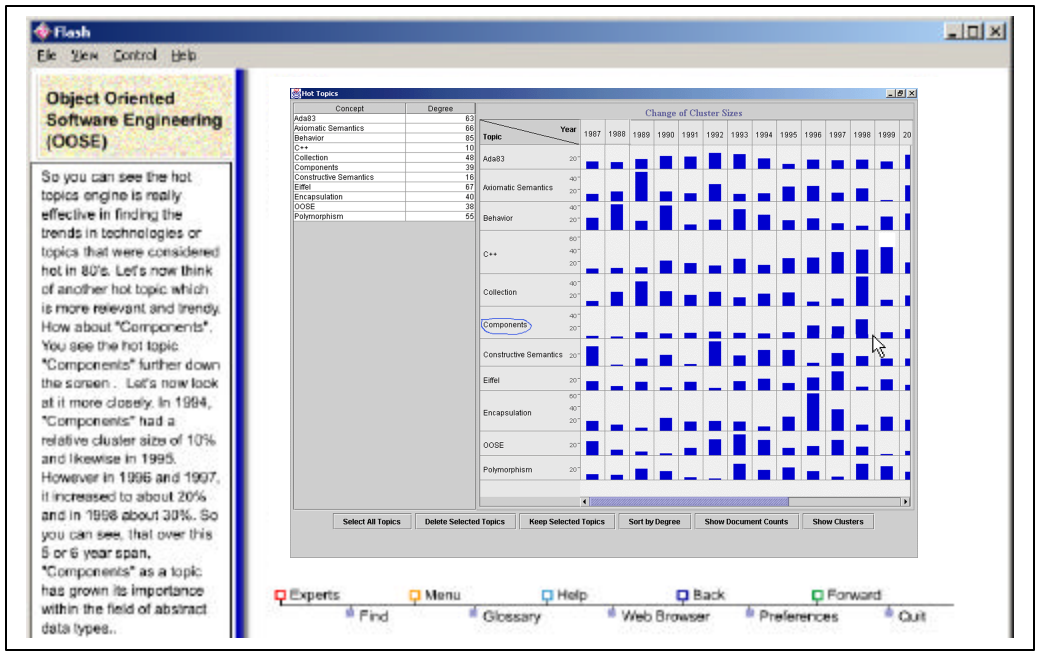

Figure 2: Screen shot of the Hot Topics Textual Data Mining System

In our research efforts to provide a network infrastructure and interface to seamlessly students looking for help, we have designed an interface and protocol for the tasks of making a connection with instructors, teaching assistants, reference librarians, and other students, and have implemented a few pieces of the design. We anticipate having a working prototype by the end of this summer, in time for user trials with students taking Fall courses.

In our design, we have identified methods for communication: by email, through an electronic chat session, though audio, by demonstration through the computer interface, and by video. After reviewing our design with Sharon Siegler, a reference librarian at Lehigh, who was skeptical of the value of a video connection, we decided to focus on text and audio-based media.

The most innovative component of the network system is the ability to demonstrate the use of computer tools for performing course research over the network. For instance, a reference librarian will be able to point to text, buttons, and type-in boxes as she is discussing their use (by audio, chat, or other communication medium). To provide a working prototype of this feature, we have chosen to develop a "window-less" plug-in to Netscape Navigator™, which we will attach to web pages being used by students. The web browser provides a common interface to external resources that is commonly used for reference searches and course work. To date, we have accomplished the following:

- Completed an architecture, interface and network protocol design for a distributed system of students, instructors, teaching assistants, and reference librarians. In our design, a central server maintains the availability and communication capabilities of experts that are communicated to the students through an application integrated with the course

multimedia system. This application provides the starting point for students searching for real-time help, or for a "show-me" demonstration that can be replayed to provide help for common questions. The detailed specifications are available through the project web site (http://www.cse.lehigh.edu/~cimel/documentation.html).

- Developed a "window-less" Netscape plug-in that can draw graphical annotations on top of any web page to which that plug-in has been attached. We plan to use a proxy server or a CGI script that modifies web pages as they are visited to include our plug-in.
- Developed a relay network application that the Netscape plug-in can communicate with through a TCP socket, and which can communicate with the network system server and relay applications of other users.

For the complete prototype, we plan to implement a peer-to-peer network topology, with the central server simply providing current availability, capability, and contact information that each relay application can use to contact the appropriate peer. Experts (instructors, TA's, or reference librarians) will make themselves accessible through the relay application that registers their availability, capability, and network address with the central server. Students will use their relay application to determine who is available to answer their questions, and to initiate a connection. We will provide many-to-one communication, where students can join in a session between a student and an expert, if the topic is of interest and given permission to do so by the expert.

Prototype evaluation and further development

In February 2001, we conducted a web-based survey, then convened a review panel to assess our plans, design documents and the prototype. Survey participants included students in the PI's senior-level class on Multimedia Development and Design, plus faculty experts. Survey results are available at http://www.cse.lehigh.edu/~cimel/eval/ReviewPanel/Feb01Results.htm and an analysis in http://www.cse.lehigh.edu/~cimel/eval/ReviewPanel/SurveyWriteup.htm. Document http://www.cse.lehigh.edu/~cimel/eval/ReviewPanel/ReviewPanelNotes.htm summarizes the recommendations of the review panel.

A second research thrust involves the ongoing development of techniques for detecting emerging topics in textual data. Major findings are reported in "Detecting Emerging Concepts in Textual Data Mining" by William M. Pottenger and Ting-hao Yang in *Computational Information Retrieval*, Michael Berry, Ed., Kluwer Academic Publishers, August 2001. In brief, we have achieved levels of precision that show an improvement in performance of six to seven times the baseline precision.

We began our user profiling research by examining server web logs in order to profile individual users[1]. The logs, however, did not contain enough information to distinguish individual users in most cases. A second finding was that individual user access was quite sparse. Users did not seem to frequent the site for very long, and during the two week period of time we chose, few users made repeat visits to the site. In order for temporal user profiling to be effective, there must be sufficient data to characterize the user's browsing activities. The web server logs however could not provide us with this type of data because users did not frequent the site enough to yield adequate temporal data.

[1] Our logs were drawn from www.ncsa.uiuc.edu for a two-week period.

These factors compounded by the uncertainty of identifying individual users caused us to abandon these logs as a viable source of data for our research. In response, we devised an approach to track the usage of students accessing multimedia courseware.

A third research effort is the design of a collaborative user interface, which will provide a seamless network connections between students and professors, teaching assistants, reference librarians, and other students. The collaborative network interface in CIMEL project will have a Client /Server architecture, as shown in Figure 3.

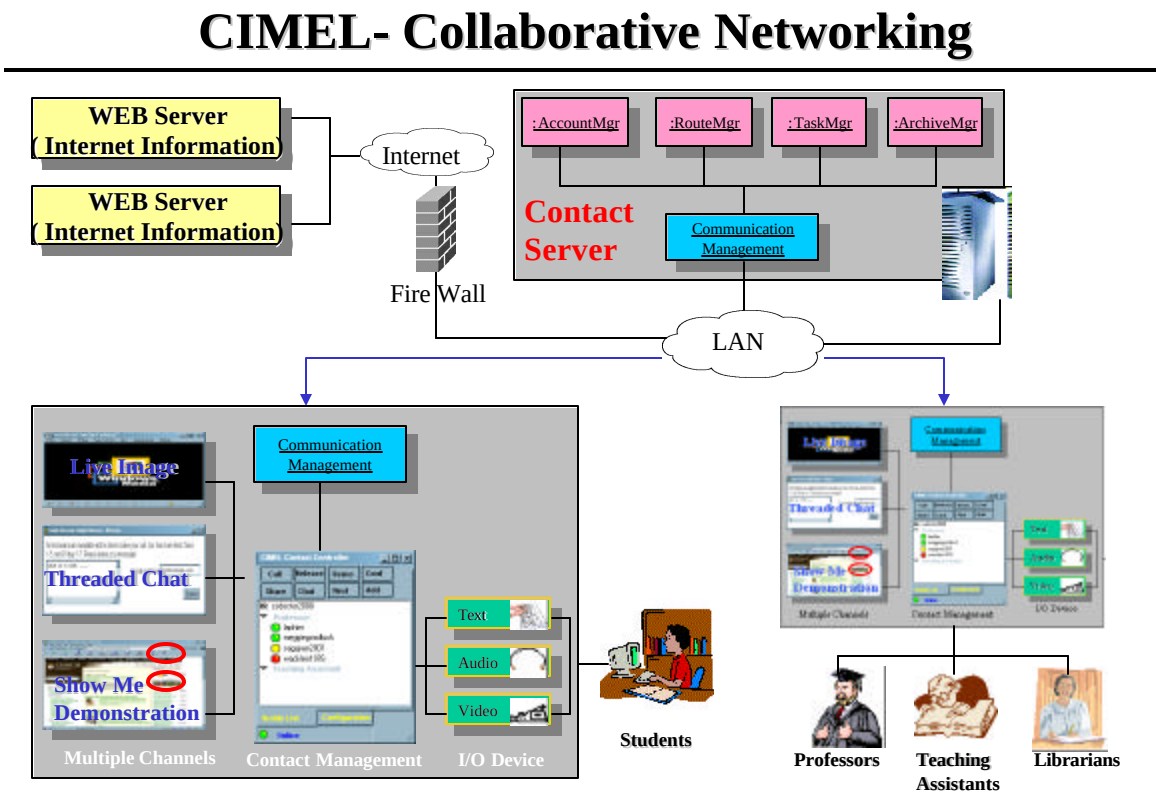

Figure 3: Architecture of CIMEL collaborative learning environment

The lower left corner of the figure shows our proposed architecture for the client site, a software system with three layers:

1. The **I/O Device** layer will manage the input/output device drivers for the collaborative network interface. I/O devices will include audio, video, keyboard and mouse.

2. The **Contact Management** layer will monitor user state, control the communication process, and let users connect to one or more other users and record their interactions.

3. The **Multiple Channels** layer will let an expert have a replica of a student's application, so that the expert can guide a student through a lesson in time, or conversely, a student can demonstrate his or her interpretation of the problem remotely. Collaborators can choose to record, annotate and archive into a demonstration database, which will effectively be a multimedia FAQ.

The upper right corner shows the server component, or **Contact Server**, which has four main modules. Account management will manage the account information of all users, including login-logout procedures. Based on current active user-list, user accounts and access history, route management will use a skill route and intelligent route algorithms to compute the call route to locate requested collaborators. Task management will schedule such tasks as outbound call, automatic email replay tasks, and cyclic management tasks etc. Archive management will provide a way to query, select, playback and update archived information.

Much of the multimedia content for this project will be derived from that of *The Universal Machine: A Multimedia Introduction to Computing*, by Glenn D. Blank and Robert F. Barnes, (WCB-McGraw-Hill, 1998), a textbook introducing Computer Science to first year students. With the help of Prof. Edwin Kay, Prof. Blank will be rewriting the textbook itself, starting this summer. The new title will introduce Java instead of C++. The associated multimedia will use the new framework developed for this project, including content introducing Java. This summer we will be redesigning the user interface and creating content introducing programming in Java.

This fall, Prof. Blank will teach both the introductory and graduate level courses, using the alpha version of the multimedia software, including alpha versions of the collaborative environment and "emerging trends" textual data-mining tool. We are currently developing a general, template-based framework for content development, using Smart Clips and the Swift dynamic content engine. This summer, they will be using this framework to develop content for the OOSE course. We are also developing evaluation tools, for tracking the use of the software, as well as pre-tests, post-tests, surveys and focus groups.

The CIMEL project web site, http://www.cse.lehigh.edu/~cimel/, includes links to a summary of project goals, all documentation and papers associated with the project, a list of team members and the project prototype, implemented in Macromedia Flash 5.

A beta version will be available for use and evaluation at Lehigh University and other campuses by the Fall of 2002. Parties interested in using or evaluating CIMEL software and textbook materials are more than welcome to contact us at glenn.blank@lehigh.edu.

The three principal investigators wish to thank God for inspiring us to begin this project and blessing the work of our hands.