BUTTERWORTH HEINEMANN

Chemical collaboratories using World-Wide Web servers and EyeChem-based viewers

Omer Casher and Henry S. Rzepa

Department of Chemistry, Imperial College, London, England

We present a "proof-of-concept" model of an Internet-based chemical collaboratory. This is based on an integration of a World-Wide Web server running the HTTP protocol, hypertext-markup language-based browsers, molecular visualizers based on Explorer EyeChem modules, and browsers implementing the virtual-reality modelling language 3D (VRML) scene description.

INTRODUCTION

A chemical collaboratory is a relatively new concept in molecular information technology, deriving from a fusion of modular visualization environments with techniques such as videoconferencing, whiteboarding, and other network-based communications techniques. In the last few years, a number of national high-speed networks based on packet switched protocols such as ATM (asynchronous transport mode) together with custom boards fitted to commercial workstations have enabled real-time molecular applications to be implemented. Our own work in this area involved the use of the UK Superjanet system running a set of Explorer-based modules that we have called EyeChem.[1] One particular EyeChem custom application called Eye2Eye enables two or more remote collaborators to export to each other a rotatable 3D image of a molecule. Each participant can manipulate the image themselves, and these operations are viewed by the other participants in real time. Such "molecular videoconferencing" can also be accompanied by a conventional audio and video link. In our case we used a commercial product called *InPerson*, which also introduces a whiteboard in which images, text, files, and other digital information can be interchanged among the collaborators.

A parallel development in the use of Internet-based networks in the last 2 years has been the remarkable growth of the World-Wide Web (WWW), a hypertext-based global information system. In the last year in particular, a number of innovative applications in the area of chemistry have been introduced,[2] including a concept termed "hyperactive" molecules,[3] in which a set of molecular coordinates stored on a WWW server can symbolically hyperlink into a WWW document. This document is retrieved from the server using a protocol known as HTTP (hypertext transfer protocol) and displayed on a so-called WWW browser using markup commands known as HTML (hypertext markup language). Using a further mechanism known as MIME,[4] which has been adapted for chemical applications,[5] the WWW browser, which is itself incapable of rendering the molecular coordinates, can pass these onto a chemical "helper" for visualization. Such a helper can be specified by the local user to be, *e.g.*, an appropriate EyeChem module, although in principle any suitable program can be selected. The logic of the communication between the various components, together with the protocols used, is shown in Figure 1.

From the aspect of designing a collaboratory involving two-way communication, the monodirectional MIME protocol operating between a document browser and the molecular visualization system is a severe limitation. In particular, it is not possible for any form of hyperlink that might be present in a molecular data file to be acted on appropriately. The molecule file is an information cul-de-sac. Whereas hyperlinks in a document composed of HTML can reference molecular data, the reverse is not true. Because HTML is basically a two-dimensional page description language and molecular data are often intrinsically three-dimensional, efficient information flow between the two different types of data file is essential. In this article we present one possible solution to this problem, and introduce an alternative and we think complementary approach of having three-dimensional molecular "scene" descriptions

Address reprint requests to Dr. Rzepa at the Department of Chemistry, Imperial College, London SW7 2AY England.
Received 25 July 1995

Presented in part at the American Chemical Society Meeting, Anaheim, California, April 7, 1995.

Journal of Molecular Graphics 13:268–270, 1995

655 Avenue of the Americas, New York, NY 10010

0263-7855/95/$10.00
SSDI 0263-7855(95)00053-9

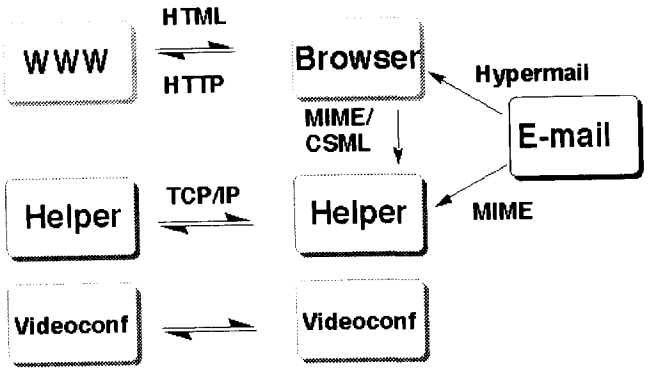

Figure 1. The interaction between WWW servers, browsers, chemical helpers, and other information systems.

defined using a newly introduced protocol known as VRML, or virtual reality modelling language.

RESULTS AND DISCUSSION

One solution to the problem of introducing communication between WWW browsers and in our case the EyeChem collection has been implemented.[6] The common client interface (CCI) is available in the form of a developers' library. Using this, we developed a new EyeChem module called EyeCCI, to achieve two-way dataflow with the NCSA Mosaic WWW browser. Thus EyeCCI can receive a so-called URL (uniform resource locator) request from Mosaic (an example might be http://www.ch.ic.ac.uk/atp.pdb) which could represent a set of molecular coordinates to be visualized by the appropriate EyeChem module. In turn, EyeCCI could initiate a request for a URL to be resolved and displayed by the Mosaic browser (say http://www.ch.ic.ac.uk/hyperactive/lprc.html). A further elaboration would be for EyeCCI to issue a request for Mosaic to retrieve a set of remotely stored molecular coordinates, and pass them back to EyeChem for visualization. In effect, we have achieved a formal implementation of the HTTP transport protocol within EyeChem by using Mosaic as a "proxy," an example of efficient symbiosis between two program systems originally designed for quite different purposes. The subtle change in the data flow compared with the initial diagram is illustrated in Figure 2.

This combination of modules means that URLs or the future URNs (uniform resource names) can now be designed into various molecular datafile formats. The one we selected to illustrate the point is the protein databank (PDB) format, which can be extended with USER commands to include URL citations. Such molecular URLs could be associated with the entire molecular datafile, with specific regions of the molecule(s), or perhaps their properties, or simply with individual atoms or bonds. Such URLs could point to other network-based resources such as other molecular datafiles, images, databases, text, or any useful and relevant collection of information. In effect, the molecular datafile becomes self-annotating, and one might even imagine that an entire scholarly paper could consist of a single molecular datafile with appropriate URLs pointing to discussion of the various features of interest. When visualized using EyeChem, resources that cannot be rendered are simply passed on to Mosaic or indeed other appropriate programs for viewing. One can even imagine a heterogeneous molecular visualization system, where for example the Explorer-based EyeChem and the AVS-based Chemistry viewer could collaborate, each processing the same data in their own way. Also enabled now are remote network-based collaborations between one user employing EyeChem and another using AVS. In effect, we are now achieving a true chemical collaboratory.

The metaphor of a collaboratory can in fact be extended further into nonmolecular "worlds." Explorer rendering is based on use of the Iris Open Inventor 3D object library and the recently proposed VRML 1.0 (virtual reality modeling language)[7] in fact uses a subset of this library. Inclusion of EyeCCI makes EyeChem therefore equivalent in functionality to a VRML browser, albeit one particularly attuned to rendering molecular 3D objects. We note with interest that several commercial VRML browsers are under development, including a product called *WebSpace* which is also based on Iris Open Inventor. This latter product can also communicate with WWW servers, via another WWW browser called Netscape. Thus the combination of EyeChem/Mosaic/CCI is partially equivalent to WebSpace/Netscape. The latter does not presently use the CCI protocol for internal communication, but an alternative. The WebSpace/Netscape product does not in fact compete directly with the EyeChem system, simply because VRML 1.0 is currently not well suited for the particular needs of representing molecular objects.[8] Thus constructs such as tapered bonds, dot surfaces, ribbon representations, etc., are beyond the current scope of VRML. We approach this problem by placing a URL pointing to the actual PDB file within a VRML file. Eye VRML will send this URL to EyeCCI. This module in turn will retrieve the file and pass it through standard EyeChem modules for handling the graphics not supported by VRML 1.0. By contrast, the WebSpace approach is for the VRML viewer to support a superset of the VRML 1.0 nodes, basically all Open Inventor nodes. The disadvantage of this approach is that any VRML file written for WebSpace might not be viewable by other VRML viewers.

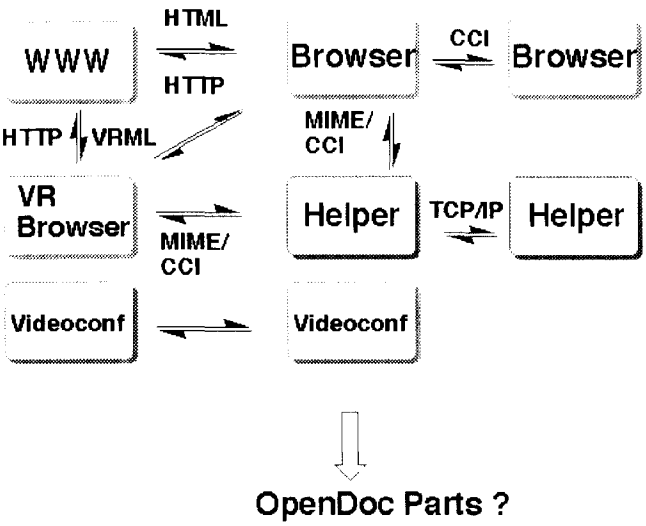

Figure 2. A WWW server/WWW browser/EyeChem/VRML interaction achieved using CCI communication.

We have used EyeChem to produce a number of molecular scenes in VRML format, available from a WWW server for immediate viewing using a suitable VRML browser.[9] The advantages of having such a combination is that one might expect the development of efficient and fast VRML browsers for a variety of computing platforms, which will achieve the true interoperability necessary for collaboratories. Another advantage of establishing links with the VRML community is to achieve continuity with other nonmolecular "scene" descriptions.

CONCLUSIONS

One might truly imagine that the scientific scholarly work of the future will consist not only of two-dimensional pages of text, images, and numerical data, but of three-dimensional navigable "worlds" of information. These worlds will have the capability of linking to other worlds, both present in archived form on a server, and interactive collaborative worlds involving real-time interaction with other scientists. In this sense, the current level of development of our EyeChem modules represents only a primitive glimpse of what might be possible in the future. The more difficult question of whether the chemical community actually wants to present results in this fashion and to use such chemical collaboratories is a more difficult one to answer.

REFERENCES

1 Casher, O., Rzepa, H.S., and Green, S. *J. Mol. Graphics* 1994, **12,** 226

2 Rzepa, H.S., Whitaker, B.J., and Winter, M.J. *J. Chem. Soc. Chem. Commun.* 1994, 1907

3 Casher, O., Chandramohan, G., Hargreaves, M., Murray-Rust, P., Sayle, R., Rzepa, H.S., and Whitaker, B.J. *J. Chem. Soc. Perkin Trans.* 1995, **2,** 7

4 Borenstein, N. and Freed, N. Internet Request for Comment (RFC) No. 1521

5 Rzepa, H.S., Murray-Rust, P., and Whitaker, B.J. Internet Draft, March–August, 1995, *rzepa-chemical-mime-types-01.txt*

6 Thompson, D. The Common Client Interface, NCSA, October, 1994. See *http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/cci-spec.html* for further details

7 Bell, G., Parisi, A., and Pesce, M. VRML 1.0; a proposed specification for a virtual reality browsing system. See the URL *http://www.eit.com/vrml/*

8 Casher, O. "A VRML Subset for Molecules," *http://www.ch.ic.ac.uk/VRML/vrmlchem.html*

9 Casher, O. and Rzepa, H.S. *http://www.ch.ic.ac.uk/VRML/*