MOLECULES TO MAPS: TOOLS FOR VISUALIZATION AND INTERACTION IN COMPUTATIONAL BIOLOGY

THOMAS E. FERRIN
Computer Graphics Laboratory
University of California
San Francisco, CA 94143-0446

EILEEN T. KRAEMER
Department of Computer Science
Washington University
St. Louis, MO 63130

Data continues to accumulate rapidly from the various genome projects and from experimental methods such as X-ray crystallography, NMR spectroscopy, and electron and confocal microscopy. The vast volume of sequence, structural, and functional data, the wide variety of analyses and annotations to be performed, and the variety of organisms and projects represented in this deluge of information combine to present the bioinformatics community with unprecedented challenges. This conference track explores challenges of enormous importance to the molecular biology and structural biology communities: tools and techniques to assist scientists in evaluating, absorbing, navigating, and correlating this sea of information, especially through visualization and user interaction.

The papers in this track describe visualization tools that span a wide range of biological data, from protein sequence, structure, and function to cellular systems. All represent powerful approaches to understanding what would otherwise be a large volume of nearly incomprehensible data. The human mind can often quickly grasp spatial relationships when presented in visual form through models, but if left in their original form, columns upon columns of numerical data will have little meaning to even the most astute scientist. The other point of interest to note is the extent to which the Internet and World Wide Web serve as the foundation for many of these visualization tools. For example, Albion Baucom and co-workers describe a web-based interactive protein modeling tool, DINAMO, that is highly portable and thus readily accessible to many users. Phillip Bourne and his colleagues describe an interactive molecular modeling environment based on the Virtual Reality Modeling Language (VRML) and designed to foster collaboration at a distance via the Internet. And Andrei Grigoriev describes a web-based browser for accessing on-line genome information services, thus providing a convenient means for a scientist to query and retrieve up-to-date information on genomic maps from a variety of sources.

Although there are several systems capable of depicting protein sequences and structures, there are far fewer systems available for describing protein function. For this reason, the graphical visualization system described by Imran Shah and Larry Hunter for exploring the predictability of enzyme function given primary protein sequence data is very interesting. Their paper describes this software tool and the types of studies it supports.

An exciting development in the field of scientific visualization in recent years has been the display of volumetric data, although often the challenge with this technique is to display massive quantities of data rapidly enough to be useful to the user. Taosong He, in his paper, describes a novel approach to rendering different levels of image quality in a rapid and controlled way using wavelet-assisted volume ray casting. He's algorithm is implemented in the program VolVis. Arie Kaufman also describes work with VolVis, including a general description of the program and its application to neurological confocal microscopy data. The successful use of VolVis to visualize calcium within nerve cells provides neurobiologists with an effective tool for investigating the function of neural systems.

Computer simulations are used in a large variety of computational modeling projects today, from computational chemistry calculations at the molecular level to the modeling of entire ecosystems. Crucial to these simulations is the accurate and facile control of the calculations. Two papers in this track address these issues. Alister Winfield describes a ''virtual'' laboratory notebook system that can be used to store the output from simulations, along with corresponding simulation input parameters, annotations by the scientist user, and a history tree and log of user interactions. C.I. Parkinson and co-workers describe a system known as MAVIS, used for visualizing and controlling computational chemistry simulations performed on a heterogeneous cluster of networked workstations. MAVIS was developed using the Application Visualization System (AVS), available on several interactive graphics workstation platforms, thus making the system quite portable. Both of these systems offer very effective approachs to the control and analysis of computer simulations.

With the advent of evermore powerful microprocessors, it is now possible to perform many of the types of graphical rendering computations previously reserved for special purpose dedicated logic on general purpose hardware. Yutaka Ueno and Kiyoshi Asai describe a portable general purpose graphics library that provides fast low-level graphics primitives on standard UNIX workstations without the need for special graphics hardware. Their library has been used in the implementation of MOSBY, a molecular structure browser which runs on non-accelerated computer hardware.

Together this collection of papers provides an excellent sample of ongoing research activities in software tool development for visualization and interaction in computational biology.

Acknowledgments

The session co-chairs are grateful to the reviewers, whose careful critiques of the submitted papers and insightful suggestions for improvements are reflected in the high quality of the resulting submissions: David Beazley, Tom Blackwell, Jerry Cox, Sean Eddy, Norberto Ezquerra, Will Gillett, Jerry Greenberg, LaDean Hillier, Conrad Huang, Philip Hubbard, Yves Jean, Eric Pettersen, Bill Ribarsky, Henry Rzepa, David Stahl, David States, and Amitabh Varshney.