AI Magazine Volume 10 Number 2 (1989) (© AAAI)

Artificial Laboratories

Mark E. Lacy

An artificial laboratory is a hypothetical computing environment of the future that would integrate mathematical and statistical tools with AI methods to assist in computer modeling and simulation. An integrated approach of this kind has great potential for accelerating the rate of scientific discovery.

0738-4602/89/$3.50 © 1989 AAAI.

One of the major characteristics of modern science is that theory and experimentation drive one another in a cyclic process of progressive refinement, leading to new conclusions about the world around us. Theory guides and directs the course of experimentation, and experimental results subsequently suggest ways in which theory must be modified. Some theories can, in fact, be discarded altogether. Over the past 30 years, computer modeling and simulation, analogous to theory and experimentation, has frequently guided scientific investigation.

John von Neumann was one of the first to pioneer and promote the use of computers to numerically study the behavior of systems and use the results as a "heuristic guide to theorizing" (Burks 1966, p. 3). The solutions provided by the computer can serve as "an aid to discovering useful concepts, broad principles, and general theories." Certainly, the use of computers in modeling and simulation of static and dynamic systems has become a significant part of scientific endeavor. Computer modeling is an extremely powerful way of working with representations that help us better understand the systems we study. We can guess that modeling and simulation will become ever more advanced in years to come, but where does the future of computers in science lie?

As rapidly as computer technology is developing, extrapolations from the present are numerous and dangerous. Although we might successfully guess that parallel processing and supercomputing will be cheaper, faster, and more widely accessible, undoubtedly many aspects of future computing would amaze us if we could use a crystal ball and see them. Let us consider for a moment, though, one direction that the use of computers in science might take.

A wide variety of mathematical and statistical software tools are available today, and the list grows daily. Means for representing and storing data and scientific information are likewise increasing, from sophisticated database and knowledge base technologies to high-speed, high-resolution graphics. Imagine a computing environment that couples all of these in a user-friendly way, using the best mathematical, computing, and AI technology to allow a scientist to quickly and easily perform all types of modeling and simulation by computer to assist in research. Imagine also that this computing environment is implemented in such a way that the human-computer interface is set up to resemble the laboratory or field environment in which real data might actually be obtained. Because of the range of computing capabilities at the scientist's disposal and the ease with which these capabilities can be brought to bear on a research problem, a computing environment of this sort might be called an *artificial laboratory.*

The artificial laboratory is, therefore, a computing environment that not only simulates the laboratory environment but also allows analysis of the data. Most likely, the laboratory environment that is simulated would be one which is relatively stable or mature. Research using custom-designed instruments, for example, might make it difficult to say with confidence that the measurement portion of the scientist–instrument–object-of-study system can be assumed to not influence the data which are obtained. Cutting-edge research might require frequent changes to computer programs used for analysis. An artificial laboratory is more suited for those situations where the laboratory technology is mature enough to permit focus only on the object of study.

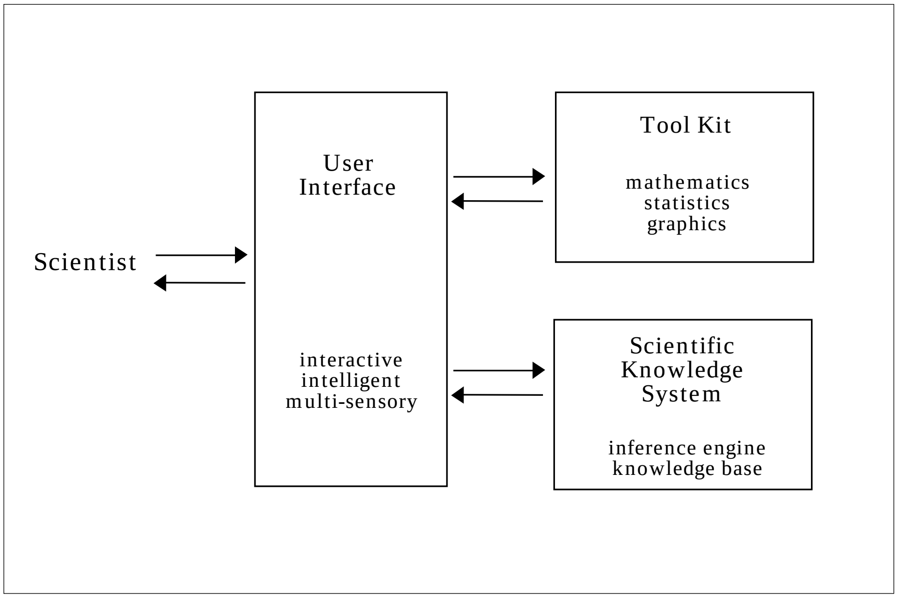


Figure 1. Primary Features of an Artificial Laboratory.

The tool kit *provides tools for computation and visualization of results. The* knowledge system *includes a knowledge base of basic scientific concepts and advice on using the tool kit and an inference engine that supports modeling. The* user interface *is a highly sophisticated subsystem which creates an artificial working environment, resembling that of an actual laboratory, for the scientist.*

The kinds of tasks that could be performed in an artificial laboratory range from testing simple thought experiments and performing classroom demonstrations of scientific concepts to finding potentially fruitful and potentially barren avenues of further investigation with the aid of the computer.

Just as forming hypotheses and testing them through experimentation are important to scientific progress, the modeling process is an important extension of the scientific method. The central paradigm in the use of artificial laboratories is that a model is another way to organize and represent knowledge. Artificial laboratories would assist a scientist in such tasks as the following:

• Developing and representing a model (structurally, graphically, mathematically)

• Testing the model (perhaps by comparing the results of simulated experiments with data from real experiments)

• Refining the model (by suggesting changes that would make the model more reasonably represent reality)

Components of an Artificial Laboratory

The three major parts of an artificial laboratory (figure 1) are (1) a mathematical and statistical tool kit and graphics procedures for displaying results; (2) a scientific knowledge system, with inference engine and knowledge base; and (3) an intelligent user interface or front end to the system.

The Tool Kit

The mathematical and statistical tool kit that would be needed for an artificial laboratory cannot simply be a lump of subroutines or procedures. Currently, for example, the IMSL Fortran subroutine library (IMSL 1987) provides a wide range of mathematical and statistical tools, but Fortran programs must be written to call these subroutines. Likewise, the mathematical and statistical procedures found in SAS (1985) software products provide the user with a host of useful tools, but some amount of SAS programming is necessary to use these tools. In the realm of tools for computational chemistry available to the public at nominal charge, the Quantum Chemistry Program Exchange provides an extensive variety of actual programs ready for use; the disadvantage with these programs is the lack of cohesiveness among the programs because they all come from different sources.

The tool kit for an artificial laboratory must be designed so that using it requires little, if any, actual programming. This effort could be approached in a variety of different ways, including using menu screens, automatic program generators, and extremely high-level programming languages. Programming as we know it, using third or fourth generation languages, would certainly inhibit the use of an artificial laboratory.

Graphics procedures must have a great deal of flexibility because scientists need to see data and information presented in many different ways. Graphs of data must allow for three dimensions, not just two, and we should be able to rotate a three-dimensional graph in any direction to get a different look at the data. Geometric diagrams of all kinds should also be possible in three dimensions, with dynamic control over the viewing possible. As with the calculations tool kit, graphics procedures should be so easy to use that they do not require extensive setup and experimentation to get something to look right. A scientist's time should be spent on scientific problems, not figuring out how to get the computer to make something work. An enormous amount of work is currently under way to make visualization in scientific computing both powerful and convenient (McCormick, DeFanti, and Brown 1987; Wolff 1988). Why is visualization so important? "The goal of visualization is to leverage existing scientific methods by providing new scientific insight through visual methods. . . . Researchers want to *steer* calculations in close-to-real time; they want to be able to change parameters, resolution, or representation, and see the effects. They want to drive the scientific discovery process; they want to *interact* with their data." (McCormick, DeFanti, and Brown 1987, pp. 3, 5).

The Knowledge System

Artificial laboratories would be designed around the particular field and applications in which they would be used. A scientific knowledge base of key concepts important to a particular field would be a critical part of any artificial laboratory. For an artificial physics laboratory, the knowledge base would include basic laws of physics, general physical principles, and key mathematical equations. For an artificial chemistry laboratory, the knowledge base would include basic concepts about chemical interactions, kinetics and thermodynamics, the periodic table, and so on. Knowledge bases for any laboratory might also include concepts from general systems theory (Lacy 1986) regarding types of models and interactions between physical units.

The knowledge base must also contain information regarding the appropriate use of tools in the tool kit. For example, the knowledge base for an artificial chemical laboratory should include a recommendation to use Gear's method for solving a system of differential equations if a kinetic model uses rate constants that span large differences in magnitude. Progress in research on expert systems that assist in simulation was reported at a recent conference.[1]

A scientific knowledge base would not be useful without an inference engine, however, to make use of the available knowledge. If the artificial laboratory maintains a database or knowledge base of models and the results obtained with each, including how the results were tested and what the outcome of the tests was, the system might be designed to suggest modifications to test. The artificial laboratory might, therefore, serve as a valuable scientific assistant.

An intelligent physiologic modeling system used for educational purposes was developed by Robert Kunstaetter (1987) at the Massachusetts Institute of Technology Laboratory for Computer Science. This system is frame based and has a deep knowledge of respiratory physiology. The admittedly gross simplification of physiological concepts embodied in the system limits its usefulness to helping medical students learn qualitative relationships revealed by perturbing various models. Nevertheless, this system is an impressive example of coupling AI technology with modeling and simulation.

Choosing the right level of abstraction would be important in designing an artificial laboratory. A high level of abstraction might be necessary for most artificial laboratories because working with abstraction on the level of first principles can be cumbersome, or the first principles might not be clearly understood or well defined. Generally, as one moves up the hierarchy from physics to chemistry and biology, for example, one moves further from wanting or being able to use the most fundamental physical principles.

It might be possible to establish a framework for developing the proper level of abstraction using general systems theory. Systems theory has been shown to provide a conceptual framework for computational chemistry (Lacy 1986). Similar frameworks could be conceived for other fields. Caution is required, however, because too high a level of abstraction would lead to an artificial laboratory so generalized that its use in particular fields would require extensive customization. A balance is necessary between too high a level of abstraction and too low a level.

Finding the proper level of abstraction might also be possible using developments in the field of qualitative reasoning and the closely related field of model-based reasoning. *Qualitative reasoning* involves the use of models to study a problem without requiring quantization of the parameters defining the model. Forbus (1988) argues that qualitative physics is central to intelligent computer-aided engineering "because it helps capture the commonsense understanding of the world that is the foundation of engineering knowledge" (Forbus 1988, p. 27) Could qualitative physics provide the level of abstraction for an artificial physics or chemistry laboratory? Could qualitative chemistry provide the level of abstraction for an artificial chemistry or biology laboratory? The tool kit for an artificial laboratory must include libraries of numeric and statistical tools. If qualitative reasoning as an approach has utility in providing the right level of abstraction for an artificial laboratory, libraries of appropriate qualitative models would also form part of the knowledge system.

The User Interface

I have described the importance of the tool kit and the knowledge base and inference engine. The usefulness of an artificial laboratory would greatly depend as well on an intelligent interface between the scientist and the laboratory. The interaction that occurs at this interface must be rapid and direct and easy to understand. The scientist should not be required to do computer programming. The scientist should not have to spend time wading through user manuals and lists of error codes.

Graphic implementations of the human-computer interface could be one of the most important advances in computer science. Today, we have many icon-based implementations that facilitate the use of a computer. For example, it is not difficult to imagine moving test tube icons around on a screen to direct an artificial chemical laboratory to simulate a reaction between two chemical species. In a generalized scheme for modeling a wide variety of systems (physical, biological, or socioeconomic), the STELLA computer modeling package uses icons for stocks, flows between stocks, factors influencing flows, and logical connectors linking stocks with factors or flows (Rowe 1988).

The direction in which development of human-computer interfaces for an artificial laboratory might follow might be similar to that pursued by Randall B. Smith (1986, 1987) at Xerox Palo Alto Research Center. Smith developed a system called the alternate reality kit (ARK). ARK is an object-oriented environment based on SMALLTALK-80 that can be used to create animated interactive systems. Everything used in ARK is represented as an actual physical object. In a simulation of simple physical concepts, for example, laws of physics

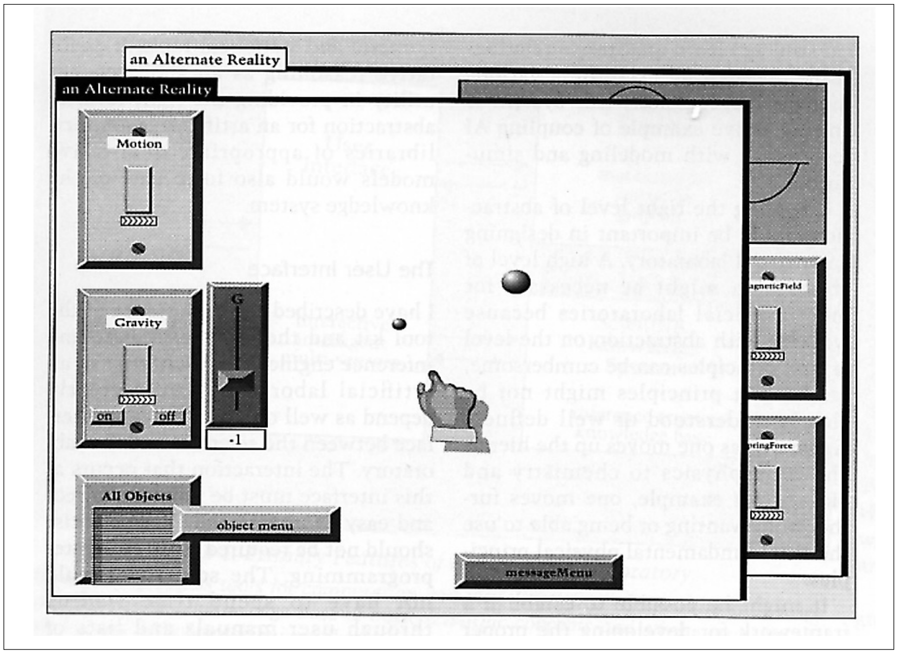


Figure 2. Typical Screen Display from the ARK System.

ARK provides a unique interface with the user. Abstractions such as the laws of nature are represented as physical objects, thereby enabling a hands-on, exploratory interaction between the user and the computer. The mouse-controlled hand icon is used to pick up, move, or throw any object on the screen. In this screen, the law of gravity is represented as an actual switch that can be turned on or off or adjusted to a desired value; the user in this image is trying to throw a moon into orbit around a planet.

Adapted from a Photo by Randall B. Smith, Xerox Palo Alto Research Center.

such as gravity and motion are turned on or off with switches shown on the computer screen (figure 2). The utility of ARK lies both in its potential for teaching through simulation and in its use as a user-interface kit.

The development of human-computer interfaces known as artificial realities (Foley 1987, National Aeronautics and Space Administration 1988) might have a profound impact on the development of an artificial laboratory. The virtual workstation concept developed by National Aeronautics and Space Administration engineer Michael McGreevy uses three-dimensional graphics and sound to present artificial realities. A head-mounted monitor that covers both eyes presents images on small screens which are viewed stereoscopically with special optics; a fiber optic glove (DataGlove, which was developed and is manufactured by VPL Research, Inc.) that records hand and finger movements gives the user a virtual hand to interact with menus shown on the display screen (figures 3a, 3b). The many possible uses for the virtual workstation include simulation in science and engineering; applications include creating virtual wind tunnels for research involving computational fluid dynamics and exploring the structure of macromolecules by touch.

Interfacing with Real Laboratories

Artificial laboratories would augment real laboratories, not replace them. Artificial laboratories would always require input from real laboratories, namely, data against which to compare simulated results. In fact, if artificial laboratories become a reality, we might wish to directly interface with real laboratories.

A data manager and a control manager are two possible interfaces between an artificial laboratory and a real laboratory (figure 4). These interfaces manage the flow of data from the real laboratory to the artificial laboratory and control of the real laboratory by the artificial laboratory, respectively. The data manager subsystem is responsible for data acquisition and preprocessing and database management; it also permits the artificial laboratory to work with either real-time data or stored data. The control manager comes into play when it is necessary to modify experimental conditions to acquire different data. Under the direction of the artificial laboratory, the control manager facilitates control of the real laboratory by sending messages to robotic devices or scientific instruments. Intense effort now under way to develop computer-integrated laboratory automation might help the development of a control interface between an artificial laboratory and a real laboratory.

Summary

A computing environment that integrates computational tools with artificial intelligence methods could be of great assistance in computer modeling and simulation in scientific research. In particular, new developments in constructing human-computer interfaces could result in an artificial reality for this environment, thereby suggesting the name artificial laboratory. The advantages of using this environment include higher productivity and creativity by avoiding direct programming and exploring problems in a manner that more closely resembles investigation in an ordinary laboratory environment.

Artificial laboratories must be managed with responsibility, however. Simulations do not produce data; they produce numbers that can mimic real-world data in some fashion. The results of simulations must be given careful scrutiny and not be blindly accepted. However, the ease with which results from simulations can be abused is not so much a drawback of artificial laboratories as it is a point of caution in scientific modeling and simulation. In addition, although computer simulation is under serious consideration as a limited alternative to the use of animals in scientific research, artificial laboratories or computer modeling should not be used to replace routine laboratory experiments whose purpose is primarily educational (for example, in a high school chemistry class). The learning acquired through hands-on experience will probably never be perfectly duplicated by an artificial laboratory.

Laboratory investigation of scientific problems has been going on for a few hundred years. As artificial laboratories are developed, we will want to interface them with real laboratories. When this event finally happens, we will have the ultimate laboratory, a fully integrated working environment specifically designed to tackle scientific problems in a way never before experienced.

Acknowledgments

I thank Franz Dill, the editor, and the reviewers for helpful comments and suggestions.

References

Burks, A. W. 1966. Editor's Introduction. In J. von Neumann. *Theory of Self-Reproducing Automata*. Urbana, Ill.: Univ. of Illinois Press.

Foley, J. D. 1987. Interfaces for Advanced Computing. *Scientific American* 257(4): 127–135.

Forbus, K. D. 1988. Intelligent Computer-Aided Engineering. *AI Magazine* 9(3): 23–36.

IMSL. 1987. International Mathematical and Statistical Library, Version 10.0. Houston: IMSL Inc.

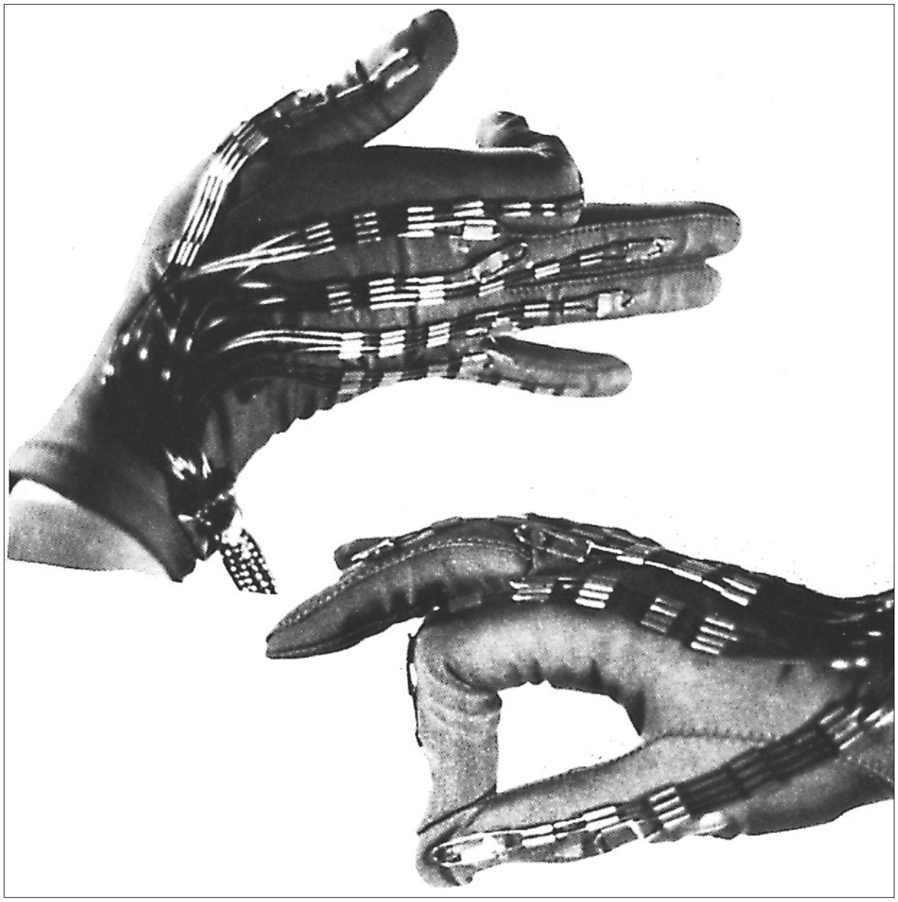
Kunstaetter, R. 1987. Intelligent Physiologic Modeling: An Application of Knowledge-Based Systems Technology to Medical Education. *Computer Methods and Programs in Biomedicine* 24: 213–225.

Lacy, M. E. 1986. Systems Theory as a Conceptual and Organizational Framework for Computational and Inferential Chemistry. *Journal of Chemical Education* 63: 392–396.

McCormick, B. H., DeFanti, T. A., and Brown, M. D., eds. 1988. Special issue on Visualization in Scientific Computing. *Computer Graphics* 21(6).

National Aeronautics and Space Administration. 1988. NASA's Virtual Workstation: Using Computers to Alter Reality. *NASA Tech Briefs:* July–August, 20-21.

Rowe, A. J. 1988. Model Maker. *Nature* 333: 608–609.



Figures 3a and 3b. Fiber Optic Glove for Human-Computer Interface.

Two views of the DataGlove, developed and manufactured by VPL Research, Inc. of Redwood City, California. This device has been used by the National Aeronautics and Space Administration as part of a virtual workstation; a head-mounted monitor presents images that can be manipulated by hand movements sensed by the DataGlove.
Photos Courtesy of Hal Elgie, VPL Research, Inc.

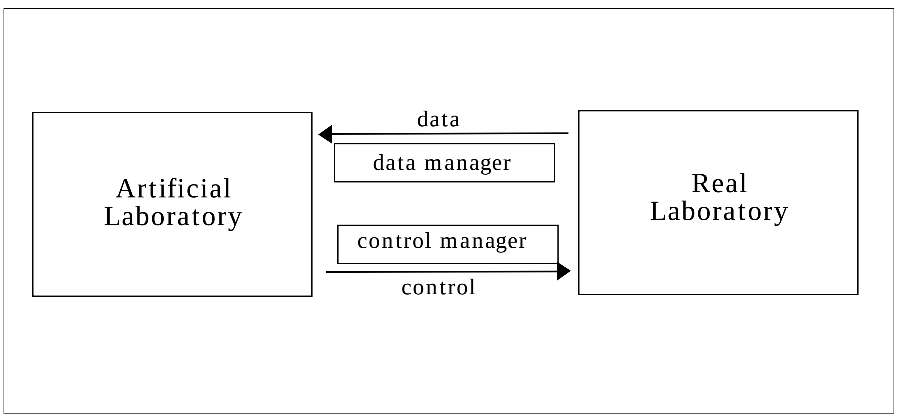


Figure 4. Possible Interfaces between an Artificial Laboratory and a Real Laboratory.

The data manager *facilitates the flow of data from the real laboratory to the artificial laboratory and is responsible for data acquisition and preprocessing and database management. The* control manager *facilitates control of the real laboratory by the artificial laboratory and is responsible for modifying experiments through connections to robotic devices and scientific instruments.*

SAS. 1985. Statistical Analysis System, Version 5. Cary, N.C.: SAS Institute Inc.

Smith, R. B. 1987. Experiences with the Alternate Reality Kit: An Example of the Tension between Literalism and Magic. Preprint. Palo Alto, Calif.: Xerox Palo Alto Research Center.

Smith, R. B. 1986. The Alternate Reality Kit: An Animated Environment for Creating Interactive Simulations. In Proceedings of the 1986 Institute for Electronics and Electrical Engineers Computer Society Workshop on Visual Language, 99–106. Washington, D.C.: Institute for Electronics and Electrical Engineers.

Wolff, R. S. 1988. The Visualization Challenge in the Physical Sciences. *Computers in Science* 2(1): 16–25.

Note

1. Conference on Expert Systems for Numerical Computing, 5–7 December 1988. Sponsored by the International Association for Mathematics and Computers in Simulation. To appear in a forthcoming special issue of *Mathematics and Computers in Simulation*.

Mark E. Lacy is responsible for biomathematics and systems research at Norwich Eaton Pharmaceuticals, Inc., a subsidiary of Procter & Gamble, P.O. Box 191, Bldg. 42, Norwich, NY 13815. His chief interests are general systems theory and applying new computational methods to chemical and biological problems in pharmaceutical research. He is a member of the AAAI, the American Association for the Advancement of Science, the American Chemical Society, the Association for Computing Machinery, the International Neural Network Society, and Sigma Xi.

Dr. Lacy chaired the symposium "Chemical Applications of Neural Network Software," presented at the 197th National Meeting of the American Chemical Society this year.