

BIOLOGICAL SYSTEM SIMULATION : A LITERATURE SURVEY

SEONHO KIM

Department of Computer Science
Virginia Tech
Blacksburg, Virginia 24061, USA

shk@vt.edu

May 1, 2003

Abstract

There are many applications of Modeling & Simulation techniques in biological field. In this paper, computer simulations in biological field are classified by the range of concern of simulations. Molecular and chemical levels of biological simulations are classified to bio-molecular simulation. This simulations deal with the dynamics of molecular, DNA, RNA, Protein and neuron cells. Metabolic system levels of biological simulations are grouped as metabolic simulation. Circulatory System Simulation, Respiratory System Simulation, Tissue and Muscle System Simulation, Immunization System Simulation, and Brain System Simulation are included in Metabolic Simulation. More wide view of simulation is sorted to Ecosystem Simulation. Population Simulation, Modeling & Simulation in Epidemiology and Agricultural System Simulation are included in Ecosystem Simulation. This paper presents brief descriptions for each simulation systems and various applications of Modeling & Simulation techniques in Biology.

Key words: Biology, simulation, modeling, DNA, molecular, computational biology, biotechnology

Table of Contents

Abstract	ii
1. Introduction	1
2. Modeling & Simulation and Biology.....	1
3. Bio-molecular Simulation.....	2
3.1 DNA Modeling	2
3.2 Protein Modeling.....	2
3.3 Neuron Modeling	3
4. Metabolic Simulation.....	3
4.1 Computational Modeling of Cellular Systems	3
4.2 Computational Modeling of Cardiovascular System	4
4.3 Computational Modeling of Neural System.....	6
4.4 Computational Modeling of Brain System.....	7
4.5 Computational Modeling of Immune System	7
5. Ecosystem Simulation	8
5.1 Population Simulation.....	8
5.2 Modeling and Simulation in Epidemiology	9
5.3 Agricultural System Simulation	10
6. Artificial Life.....	11
6.1 Emergent Properties	11
6.2 Evolutionary Computation	11
6.3 Genetic Algorithm.....	12
7. Visualization	12
8. Drug Design.....	13
9. Conclusions	14
References	15

1. Introduction

Life scientists in the future may be characterized more by their ability to use the computer for data analysis and simulation than by their ability to use the traditional microscope. This is so because biology has a greater potential for mathematical development than other sciences. Living systems are made up of interacting chemical and physical processes, all of which can be described in mathematical terms. However, these systems are so immensely complex that they have resisted mathematical analysis by the classical methods successfully used by chemists and physicists. Only within the last four decades have biologists been able to use digital computers to begin to deal with these complex systems. Simulation studies pervade the recent literature of almost every biological field. Biologists with simulation skills find them to be useful in comprehending biological models quickly, separating sense from nonsense. Simulation techniques ought to be as familiar to practicing biologists as are data analytical techniques, which have been part of biological training for decades.

2. Modeling & Simulation and Biology

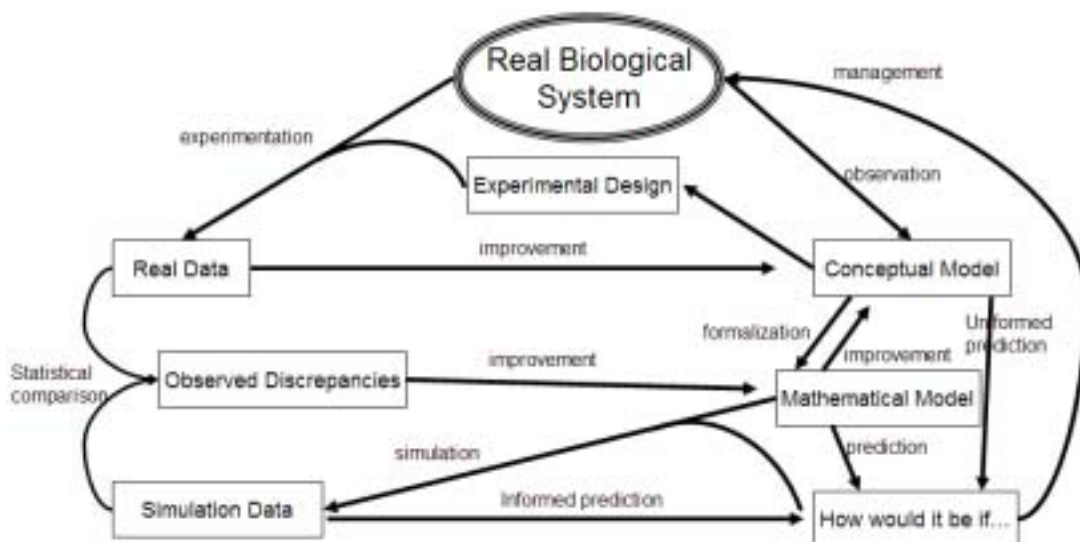


Figure 1 : Role of Modeling and Simulation within the process of research in Biology[Keen 1992]

Figure 1 is showing the role of M&S in the process of biology research. A mathematical model results from a formalization of the conceptual model, usually as an equation describing the response of a system to some variable such as time or temperature. The mathematical model may also be derived from statistical analysis of the experimental data, or by a combination of these two approaches. In any case, the thinking required to formulate the mathematical model generally improves the conceptual model. This improved model will in turn suggest further experimentation. Repeated excursions through the loop in Figure 1 can result in further improvements to both the mathematical model and conceptual model.

3. Bio-molecular Simulation

Bio-molecules are complex dynamic systems: the analysis and interpretation of their motion is important for understanding their function. Computer simulation plays a key role, providing the microscopic details of bio-molecular dynamics and characterizing the spectrum of temporal and spatial correlations in the atomic motions e.g., local sidechain isomerisations, secondary structure shifts and global domain motions.

DNA, RNA, Protein, Neuron modeling are included in this fields.

3.1 DNA Modeling

DNA provides the basic blueprints that outline the basic building blocks of the cell. It is composed of four basic molecules, called bases, which make up a coding system that outline how proteins are made. Simulation techniques are used for understanding DNA molecule, DNA topology, DNA structures and functions.

3.2 Protein Modeling

New genomic technology has produced an explosion of new protein sequence data. In principle, the proteins coded by these sequences represent thousands of new targets for pharmaceutical design. In practice, the biological function of these proteins must be determined to assess their suitability as a target, and structures ascertained to employ structure based drug design methods to design new pharmaceutically active compounds. While advances in experimental techniques such as high throughput crystallography are substantially accelerating the rate at which protein

structures are being determined, this rate is still significantly slower than the production of new sequence information. Consequently, computational methods to determine structure and/or function are highly desirable. Simulation techniques are used for sequence comparison, sequence annotation, structure analysis and structure prediction of protein.

3.3 Neuron Modeling

The growing interest in neuron modeling parallels the increasing experimental evidence that the nervous system is extremely complex. In fact, modeling is as essential as laboratory experimentation to the understanding of structure-function relationships in the brain.

4. Metabolic Simulation

Metabolic system levels of biological simulations are grouped as metabolic simulation. Circulatory System Simulation, Respiratory System Simulation, Tissue and Muscle System Simulation, Immunization System Simulation, and Brain System Simulation are included in Metabolic Simulation.

4.1 Computational Modeling of Cellular Systems

Research in this broad area covers several levels of biological organization and applies to every organ and system in health and disease. The goal is to achieve a quantitative understanding of cell function, including genetic control, gene expression, molecular interactions within the cell, intracellular signaling, cell mechanics and motility, and cell proliferation. Research efforts are directed at achieving an integrated computational model of the cell based on studies of molecular-based cell mechanics, intracellular transport, kinetics of ion channels, transporters and receptors, and cell interactions with the environment. The work provides the foundation for building computational models of different cell types. Gene and drug delivery, development and testing of new drugs and therapies, and cell and tissue engineering all stand to benefit from the application of these computational models. With the human genome sequence on the horizon, the problems of gene interactions, and control of genetic circuits and gene expression, combine to become one of the grand challenges of bioengineering. Such complex genetic systems require the application of state-of-the-art engineering methodologies to characterize function. This area will

be a focus of Center research. Molecular-based modeling of cellular structures represents the cutting edge in cell engineering research. This approach is capable of predicting mechanical, electrical, and transport characteristics of the cell in terms of the molecular biophysical properties of its constituents (e.g., protein-protein interaction, protein and lipid diffusion within the plasma membrane, receptor-ligand and ion channel kinetics). Several laboratories pursue research in diverse areas that include electromechanical and metabolic properties of cardiac myocytes, mechanics and motility of non-muscle cells, cell-cell and cell-polymer surface interactions in applications to tissue-engineered constructs, and drug and gene delivery to the cell. For computational models to become increasingly realistic, they must be based on structural and physiological data. At present, there have been few efforts to organize such information beyond the level of genomic and proteomic databases.

4.2 Computational Modeling of Cardiovascular System

Some computational biologists engage in basic and clinical research in computational modeling of cardiovascular function. The objectives of their research are to understand the fundamental biochemical, biophysical, and mechanical mechanisms governing electro-mechanical function of the normal heart; understand the genetic and molecular origins of heart disease and their relationships to cardiac function at the cellular and whole-heart level; understand the electrical and mechanical properties of blood flow in large and small vessels, the mechanisms of oxygen delivery to tissue, and the nature of blood cell-surface interactions in thrombus formation; and develop theoretical and computational approaches for the design of novel cardiovascular therapeutics. Research spans levels of analysis from that of the gene to the organ. At the genetic level, researchers seek to identify genes whose expression is altered in cardiovascular disease. Investigators then use computational models to examine the relationship between altered patterns of gene expression and biochemical, biophysical, and mechanical functioning of cells and tissue. A particular interest is to determine how altered patterns of gene expression may give rise to potentially fatal arrhythmias within the heart. At the molecular level, researchers use heterologous expression systems, site-directed mutagenesis, and patch-clamp/whole-cell recording methods in structure-function studies of cardiac ion channels. The data is being used to develop new computational models of channel behavior and drug-channel interactions.

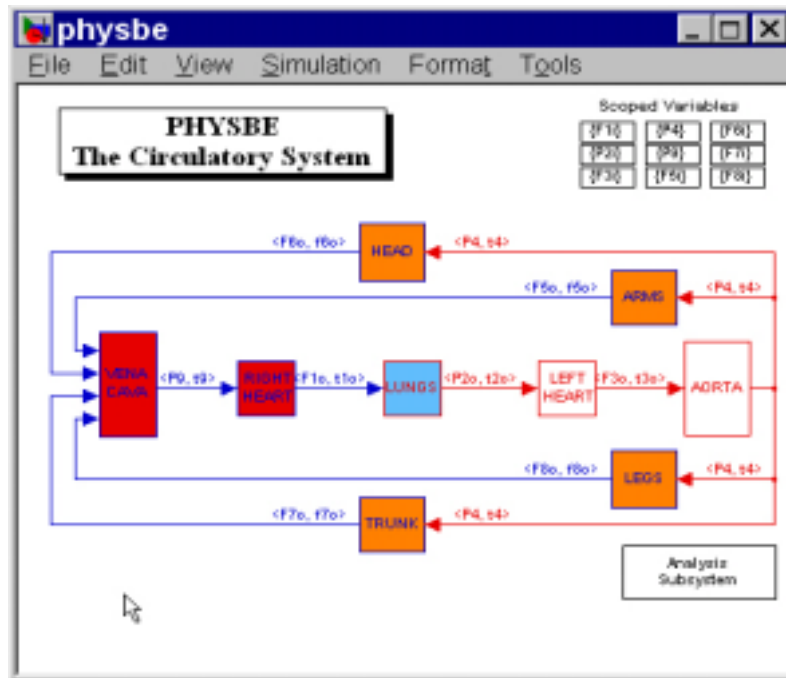


Figure 2 : Visual Simulation Environment of PHYSBE Circulatory System

At the cellular level, researchers rely on genetic and pharmacological techniques to probe the contributions of specific ion channels to cardiac excitability and recovery. With fluorescence-based imaging of intracellular calcium, scientists study mechanisms of intracellular calcium entry, cycling, and release from sarcoplasmic reticulum in both health and disease. Researchers apply data from these experiments to develop integrative models of the cardiac cell-models that include descriptions of intracellular biochemical pathways, membrane biophysics, intracellular calcium handling mechanisms, and mechanical force generation. These models are being used in the "numerical screening" of anti-arrhythmic drugs. Researchers are also constructing three-dimensional computational models of the heart to investigate the basis of arrhythmia in heart disease. Biophysically detailed models of the microvasculature are also being designed and applied to the analysis of electro-mechanical control of blood flow, and oxygen delivery to tissue.

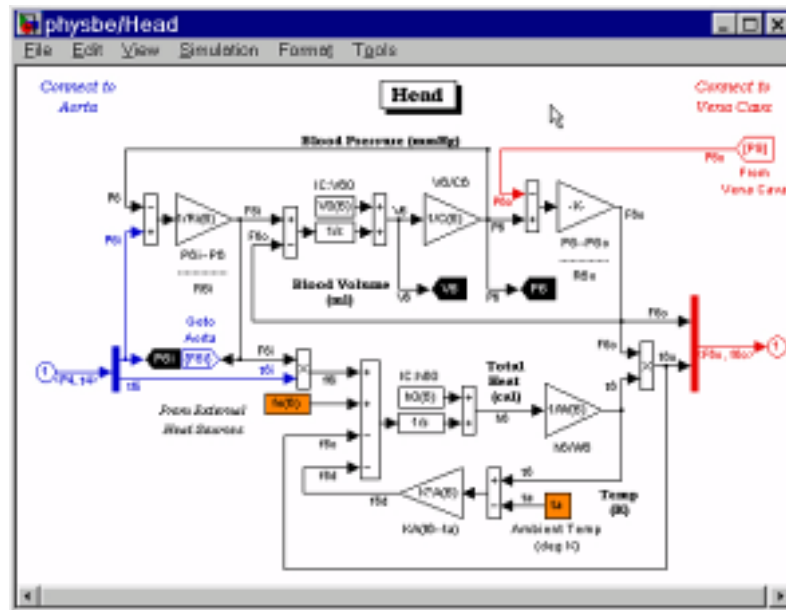


Figure 3 : Detail view of Head Section in PHYSBE Circulatory System

Figure 2 shows PHYSBE Circulatory System made by MATLAB co. PHYSBE models blood flow as a transport mechanism and can be used to simulate the flow of oxygen, nutrients, heat or chemical tracers in bloodstream. The body's circulatory system can be divided into three distinct subsystem: the lungs (pulmonary circulation), the heart (coronary circulation), and "everything else" (systemic circulation). When user click "head" in this diagram, it shows detailed sub-simulation system of head as Figure 3 shown.

4.3 Computational Modeling of Neural System

One major part of computational biology is modeling neural information processing. Activities focus on two major areas: computational modeling of the encoding and processing of complex acoustic stimulus in the auditory periphery and central nervous system and computational modeling of processes involved in motor learning and control in the brain. Scientists in the Neural Encoding Laboratory conduct research on the representation and processing of complex acoustic stimulus. General research questions include how acoustic information is encoded in neuronal spike discharge patterns; how central auditory nuclei are organized to extract information about sound from input neural activity; and how best to represent the information processing activities of neural systems in computational models. Investigators are particularly

interested in understanding and modeling the synaptic organization and acoustic processing within the dorsal cochlear nucleus and auditory cortical fields. Modeling techniques include signal detection and information theoretic analyses of neural spike trains, biophysically detailed modeling of specific auditory neurons, and neural network modeling of synaptic interactions. Researchers in this field use robotics, computational neuroscience, and functional brain imaging to discover the principles of learning in humans. Programmed robots produce novel force field environments with which human subjects interact. Subjects explore these novel environments by moving the robot handle. As subjects practice the movements, an internal neural representation of the novel force field is formed, and new control laws regulating limb movement are established. The objective of computational neuroscience is to understand, through combined experimentation and computational modeling, how this neural representation develops and how the brain generates new adaptive control laws regulating movement.

4.4 Computational Modeling of Brain System

The ability to simulate an entire human brain would be incredibly useful. It would give us insights into aspects of its function that are difficult to examine *in vivo*. It could help us develop an understanding of what happens when things go wrong and help us develop treatments. We could also gain an insight into how thoughts come to arise within the real thing. Such a system could form the basis for uploading minds into computers for those that believe that such things are possible. However there are many practical complications to simulating a human brain. The brain is an incredibly complicated apparatus estimated to contain around 100 billion neurons which may connect to thousands more. In a typical simplification, a neuron is taken as being a simple processing unit that takes a weighted sum of its inputs and activates accordingly. Quite large arrays of such units can be easily simulated in a computer. However, even in the presence of such simplification, simulating the complete array of neurons is well beyond the capabilities of a top of the range personal computer. Indeed, even the most sophisticated supercomputers would be strained to their limits.

4.5 Computational Modeling of Immune System

When a foreign substance introduced into human bodies, immune system acts to neutralize that substance. This is a complex process involving the collective and coordinated response of

approximately 10^{12} cells. In the effort to fit detailed experimental findings into a comprehensive model of the immune system, computer models and simulations are just beginning to play a role.

5. Ecosystem Simulation

More wide view of simulation is sorted to Ecosystem Simulation. Population Simulation, Modeling & Simulation in Epidemiology and Agricultural System Simulation are included in Ecosystem Simulation.

5.1 Population Simulation

The prediction of population trend of some species, such as endangered species, harmful species to human or crop and agricultural species is helpful to human to make some critical decisions. Population models are an effective method to better understand the factors that might cause population decline or ascent. Before the computer simulation in this field, mathematical models have played the role in many areas of population biology but empirical researchers in biology were finding that real populations often exhibited complex structures and processes that could not be represented adequately in rigorous mathematical models. For example, classical mathematical population genetic models often assumed for analytical convenience that regional populations were subdivided into permanent local subpopulations of fixed size, with essentially no significant internal structure. However, field biologists often found populations of animals, such as Old World monkeys, divided into local groups that were both highly complex and dynamic. In contrast to mathematical models, data structures native to computers conveniently represent many of the complexities observed in real populations. Finally, cheap, fast, microcomputers have become available, making simulation more accessible as a routine modeling tool for biologists. There are hundreds of good examples of biological population simulation systems. For examples, RICEPSM is a software for rice population simulation, WebBeePop is a web service for honey bee population simulation, ATDSIM is an american dog tick population simulation system and MALSIM is a population simulation system for Anopheline mosquitoes which transmit malaria virus.

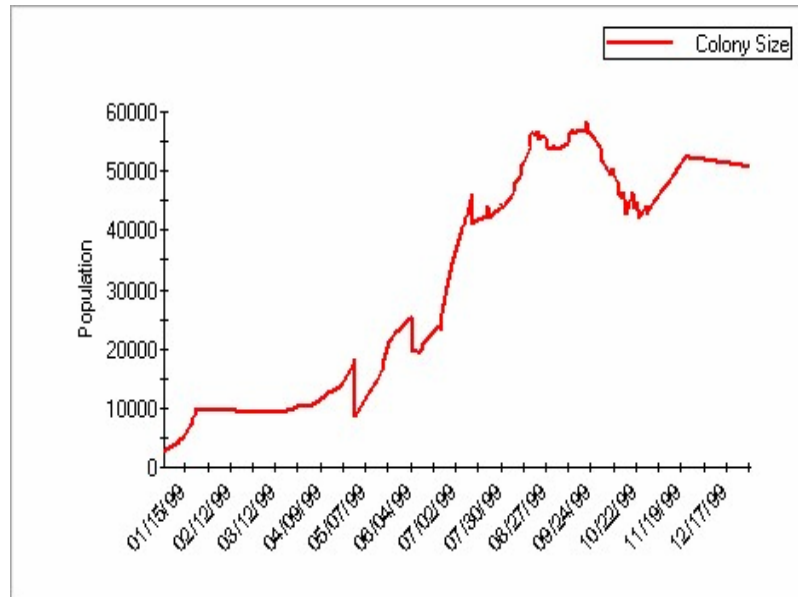


Figure 4 : Simulation result of WebBeePop

Figure 4 shows a population simulation result of WebBeePop when initial colony number was 1000 and city was Atlanta. WebBeePop is a web-based simulation application. Once user have selected a city and initial population, WebBeePop will simulate the changes in the population from January 1st to December 31st. By using daily weather information, WebBeePop makes updates of the population size. The weather information it uses to run this simulation are

- Day length : because the queen lays more eggs on long days.
- Mean daily temperature : because the queen lays more eggs on warm days and foragers fly only if the temperature is sufficiently high.
- rain and wind speed : because foragers do not fly if it is windy or rainy.

5.2 Modeling and Simulation in Epidemiology

Two weather-based models for simulation of the population dynamics of disease vectors were used to assess the possible effects of climate change due to increased levels of CO₂ in the atmosphere on vector-borne disease transmission in the United States. The first model (ADTSIM) simulates population dynamics of the American dog tick (ADT), *Dermacentor variabilis*, which is the primary vector of Rocky Mountain Spotted Fever (RMSF) in the eastern U. S. This model included the effects of temperature and atmospheric moisture on the life processes of the tick. The density of adult ticks was used as an indicator of RMSF transmission

potential because this is the stage normally involved in transmission of human cases. The second model (MALSIM) simulates the population dynamics of Anopheline mosquitoes and the transmission of malaria between mosquitoes and humans. This model simulates direct incidence of malaria, assuming that the disease is reintroduced into a human population that is continuously exposed to mosquito bites. The effects of temperature, atmospheric moisture, and rainfall are included in MALSIM. The malaria vector considered in these simulations is *Anopheles quadrimaculatus*, which presently exists in many areas of the eastern half of the U. S. and was the primary vector of malaria in the south and east when the disease was endemic.

5.3 Agricultural System Simulation

Many soil, climatic, plant, and management factors affect the way a crop will respond to irrigation, fertilizer, and other management practices. Determining appropriate crop management strategies under these uncertainties has major economic and environmental implications. Computer simulation models of the soil/crop/atmosphere system can make a valuable contribution to both furthering understanding of the processes determining crop responses and predicting crop performance in different areas.

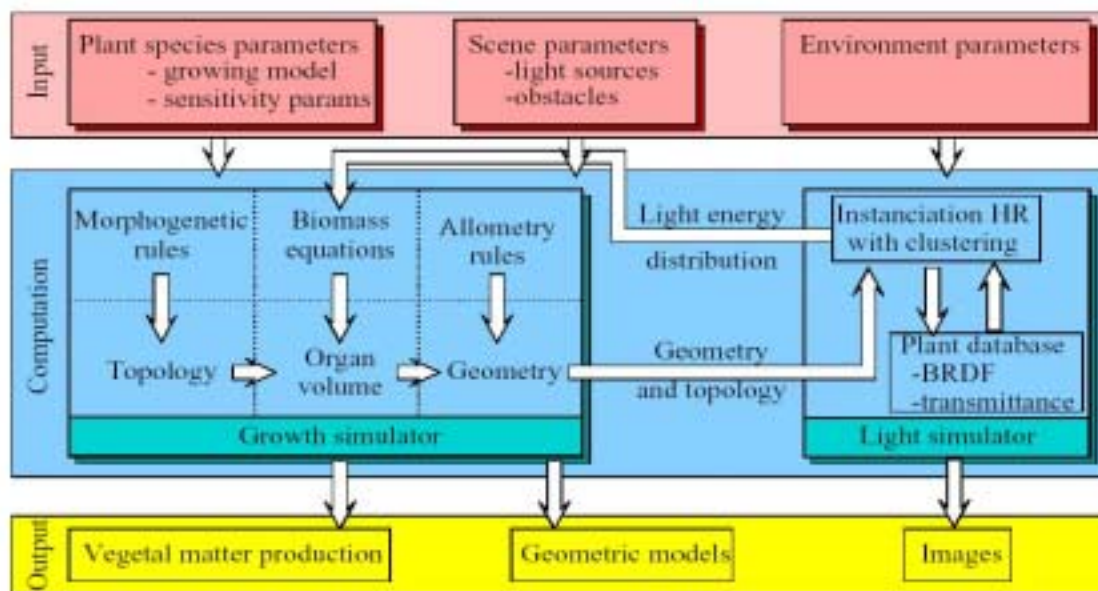


Figure 5 : Crop Growth Agricultural Simulation System

With the increasing availability of personal computers, user-oriented simulation models will greatly facilitate the task of optimizing crop and nutrient management and deriving recommendations concerning crop management. They also can be used to investigate environmental and sustainability issues of agro-ecosystems.

Figure 5 depicts the architecture of a crop growth agricultural simulation system [INRIA 2003]. The crop growth simulator kernel on the left-hand side is in charge of three tasks: computing the plant topology, computing the volume of organs and computing the geometry of the plant. In addition to the lighting simulation module, mechanical simulators to deal with internal mechanical constraints when inferring the geometry of branches are exist.

6. Artificial Life

Artificial Life is the name for a collection of various disciplines. It is about the study of non organic organisms, of life like behavior beyond the creation of nature. In the quest of behavioral biologists to understand the behaviors of nature and computer scientists to create new and better programs, paths increasingly crossed. Research of different kinds can, in many cases, lead to unifying knowledge.

Several properties of life, such as emergent properties, evolution, diversity, mutation, inheritance and reproductions are important issues in this study.

6.1 Emergent Properties

An emergent property is created when something becomes more than the sum of its parts. The most extreme example of an emergent property we know of is life on this planet. We are more than the sum of our parts. Half a human does not work without the other half, but as a whole we are capable of very complex behavior.

6.2 Evolutionary Computation

Evolutionary computation is a term used to describe computer-based problem solving systems which use computational models of some of the known mechanisms of “evolution” as key elements in their design and implementation. A variety of evolutionary computations have been proposed. The major one is: Genetic algorithm. Evolutionary computation has conceptual bases

of simulating the evolution of “individual” structures via processes of “selection”, “mutation” and “reproduction”. The processes depend on the perceived “performance” of the individual structures as defined by an “environment.”

6.3 Genetic Algorithm

The Genetic Algorithm is a model of machine learning which derives its behavior from a metaphor of some of the mechanisms of “evolution” in nature. This is done by the creation within a machine of a “population” of “individuals” represented by “chromosome”, in essence a set of character strings that are analogous to the base-4 chromosomes that we see in our own DNA. The individuals in the population then go through a process of simulated "evolution."

7. Visualization

Visualizable objects in biology and medicine extend across a vast range of scale, from individual molecules and cells, through the varieties of tissue and interstitial interfaces, to complete organs, organ systems and body parts, and include functional attributes of these systems, such as biophysical, biomechanical and physiological properties. Visualization in three dimensions of such objects and their functions is now possible with the advent of high resolution topographic scanners and imaging systems. Medical applications include accurate anatomy and function mapping, enhanced diagnosis, accurate treatment planning and rehearsal, and education/training. Biologic applications include study and analysis of structure to function relationships in individual cells and organelles. The potential for revolutionary innovation in the practice of medicine and in biologic investigations lies in direct, fully immersive, real-time multi-sensory fusion of real and virtual information data streams into online, real-time visualizations available during actual clinical procedures or biological experiments. Current high-performance computing, advanced image processing and high fidelity rendering capabilities have facilitated major progress toward realization of these goals. With these advances in hand, there are several important applications of 3-D visualization that will have a significant impact on the practice of medicine and on biological research.

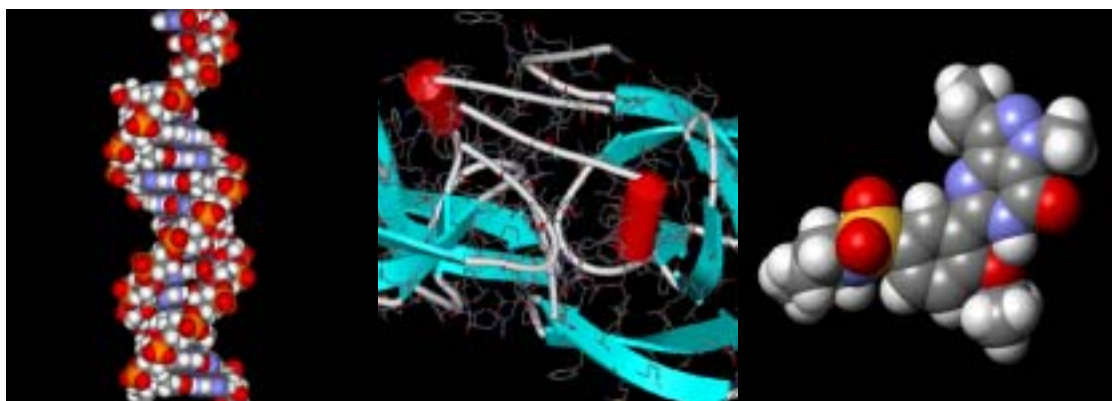


Figure 6 : Biological Data Visualization (3D DNA structure, HIV Virus, Viagra)

8. Drug Design

The input of bio-computing in drug discovery is twofold: firstly the computer may help to optimize the pharmacological profile of existing drugs by guiding the synthesis of new and "better" compounds. Secondly, as more and more structural information on possible protein targets and their biochemical role in the cell becomes available, completely new therapeutic concepts can be developed. The computer helps in both steps: to find out about possible biological functions of a protein by comparing its amino acid sequence to databases of proteins with known function, and to understand the molecular workings of a given protein structure. Understanding the biological or biochemical mechanism of a disease then often suggests the types of molecules needed for new drugs. In all cases, the aim of using the computer for drug design is to analyze the interactions between the drug and its receptor site and to "design" molecules that give an optimal fit. The central assumption is that a good fit results from structural and chemical complementarity to the target receptor. The techniques provided by computational methods include computer graphics for visualization and the methodology of theoretical chemistry. By means of quantum mechanics the structure of small molecules can be predicted to experimental accuracy. Statistical mechanics permits molecular motion and solvent effects to be incorporated. Basically statistical mechanics is a three-dimensional equivalent of describing the position of billiard balls using Newton's law of motion.

9. Conclusions

Modeling & Simulation techniques has applications in medical, pharmaceutical studies. And it is being used for epidemiological, agricultural and educational purposes also. Modeling and Simulation techniques are important tools to conduct researches for biologists. It makes biologists understand more complex processes in biological system which were not possible to explain with mathematical biological system. Modeling and Simulation techniques help us to understand biological system, to answer biological queries, to develop new medicine, to develop more strong and more productive species and to improve the environment.

References

- Keen, Robert. and Spain, James D, (1992), *Computer Simulation In Biology, A BASIC Introduction*, Wiley-liss Press, New York, NY
- Wilson, Will., (2000), *Simulating Ecological and Evolutionary System in C*, Cambridge University Press, New York, NY
- Iyengar, S. Sitharama., (1998), *Computer Modeling and Simulations of Complex Biological System*, CRC Press
- Toquenaga, Y. (1990), “A Mechanistic Model of Contest and Scramble Competition Observed in Callosobruchus Species,” *In Proceedings of the Fukuoka Symposium of Theoretical Ecology*, Aug. 31-Sep.2. Category A 14p. 46, Fukuoka, Japan
- Goldberg, D.E. (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley
- Kitano, Hiroaki., (1993), *IDEN TEKI Algorithm*, Sangyo Tosho, Tokyo
- Davis, L. (1991), *Order-based Genetic Algorithms and the Graph Coloring Problem*, *Handbook of genetic algorithms*, van Nostrand Reinhold
- Fujiki, C. (1987), “Using the Genetic Algorithm to Generate LISP Source Code to Solve the Prisoner’s Dilemma”, *In Proceedings of 2nd International Conference of Genetic Algorithms*, 1987
- Zani, P.A (1999), “The comparative evolution of lizard claw and toe morphology, clinging performance, and microhabitat use”. *Journal of Evolutionary Biology*, v13 316-325 1999
- Martins, E.P (1996), “Conducting phylogenetic comparative analyses when the phylogeny is not known” *Evolution* v50 12-22
- Martins, E.P (2002), “Adaptation and the comparative method: A computer simulation study”, *Evolution* v56 1-13
- Lengauer, T., Ed. (2002), *Bioinformatics – From Genomes to Drugs*, Volume I: Basic Technologies; Series: Methods and Principles of Medicinal Chemistry Vol. 14

- B.Kohler, R. Puzone, et al., (2000), “A systematic approach to vaccine complexity using an automaton model of the cellular and humoral immune system”, *Vaccine 19*, 862-876
- Kleinstein, Steven H. and Philip E. Seiden, (2000), “Simulating the Immune System”, *Computing in Science and Engineering*, pp69-77
- Seiden, P.E. and Celada, F.,(1997), “A Simulation of the Immune System, Experiments in machine”, in *Some New Direction in Science on Computers*, World Scientific Press, Singapore
- Haile, D.G. and Weidhaas, D.W., (1977), “Computer Simulation of Mosquito Populations (*Anopheles albimanus*) for comparing the effectiveness of control techniques”, *Journal of Medicine*, page553-567
- Allen, M.P. and Tildesley, D.J., (1987), *Computer Simulation of Liquids*, Oxford University Press, New York, NY
- Bates,A.D. and Maxwell,A., (1993), *DNA Topology in Focus series*, Oxford University Press, New York, NY
- Branden,C. and Tooze,J.,(1999), *Introduction to Protein Structure*. Second edition, Garland Publishing Inc, New York, NY
- Bratley, P., Fox, B.L. and Schrage, L.E., (1987), *A Guide to Simulation*, Springer Verlag, New York, NY
- Brooks, C.L., Karplus, M. and Pettitt, B.M.,(1988), *A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*, Wiley Interscience, New York, NY
- INRIA, (April 2003), website, <http://www.inria.fr>, France