BIOLOGICAL INFORMATION AND COGNITIVE INFORMATICS

Matthew X. He
Division of Math, Science, and Technology
Nova Southeastern University
Ft. Lauderdale, FL 33314 USA
Email: hem@nova.edu

Abstract: This paper briefly reviews the intersections and connections between these two emerging fields of bioinformatics and cognitive informatics through a systems view of emerging pattern, dissipative structure, and evolving cognition of living systems. It is hoped that this brief review encourages further exploration of our understanding of the biological knowledge of cognition, perception, learning, memory, thought, and mind.

Key-Words: Bioinformatics; cognitive informatics; patterns; data; information; knowledge discovery

1. Introduction
Bioinformatics is the comprehensive applications of mathematics, science, and a core set of problem-solving methods to the understanding of living systems. It will have profound impacts on all fields of biological and medical sciences. Cognition is viewed as a process of living systems. Cognition is an abstract property of advanced living organisms. It is studied as a direct property of a brain or of an abstract mind on sub-symbolic and symbolic levels.

Cognitive informatics studies cognition and information sciences that investigates the processes of the natural intelligence. As both fields continue their rapid development and progress, it is a central challenge to understand the biological basis of cognition, perception, learning, memory, thought, and mind.

The time seems ripe to bring these varied topics together to focus on our understanding of the emerging patterns, dissipative structures, and evolving cognition of living systems through a process of experimental application, scientific computation, and theoretical abstraction. This paper briefly reviews the intersections and connections between these two emerging fields of bioinformatics and cognitive informatics through a systems view of emerging pattern, dissipative structure, and evolving cognition of living systems.

2. Biological Data
The potential data objects in bioinformatics are illustrated below in Figure A:

FIGURE A Potential data objects in bioinformatics
The general biology-driven problems in bioinformatics include

- Find functionally significant motifs in a family of protein sequences;
- Develop techniques to detect alternate genetic codes;
- Develop techniques to identify the extent of horizontal gene and intron transfer;
- Develop techniques to help understand the role of DNA repeats in genome evolution.

Over the past few decades, major advances in the field of molecular biology, coupled with advances in genomic technologies, have led to an explosive growth in the biological data generated by the scientific community. This deluge of genomic information has, in turn, led to an absolute requirement for computerized databases to store, organize, and index the data and for specialized tools to view, analyze, and interpret the data. Bioinformatics is an emerging field of science in which biology, computer science, and information technology merge to form an emerging discipline. The ultimate goal of the field is to enable the discovery of new biological insights and hidden patterns of living systems at every level.

Given a sequence of data such as a DNA or amino acid sequence, a motif or a pattern is a repeating subsequence. Such repeated subsequences often have important biological significance and hence discovering such motifs in various biological databases turns out to be a very important problem in computational biology. Of course, in biological applications the various occurrences of a pattern in the given sequence may not be exact and hence it is important to be able to discover motifs even in the presence of small errors. Various tools are now available for carrying out automatic pattern discovery. This is usually the first step towards a more sophisticated task such as gene finding in DNA or secondary structure prediction in protein sequences at system level.

2.1 Systems Biology

Systems biology is an emergent field that aims at system-level understanding of biological systems. It focuses on systems that are composed of molecular components. This may include the understanding of structure of the system such as gene regulatory and biochemical networks and the understanding of dynamics of the system both quantitative and qualitative analysis. There are numbers of exciting and profound issues that are actively investigated, such as robustness of biological systems, network structures and dynamics, and applications to drug discovery. Systems biology and network biology are in their infancy, but these are the areas that have to be explored and the areas that demonstrate the main stream in biological sciences in this century [4.5].

2.2 Biological Networks

One of the ultimate goals of biological networks is to improve our understanding of the processes and events that lead to pathologies and diseases. The analysis of biological pathways can provide a more efficient way of browsing through biologically relevant information, and offer a quick overview of underlying biological processes. Protein interactions help put biological processes in context, allowing researchers to characterize specific pathway biology. Hence, the analysis of biological networks is crucial for the understanding of complex biological systems and diseases. The analysis of protein interaction networks is an important and very active research area in bioinformatics and computational biology [2].

3. EMERGING PATTERN, DISSIPATIVE STRUCTURE, AND EVOLVING COGNITION

Patterns, structures, and rules arise and play an important role in living systems and nearly all branches of science. This is particularly true in mathematics, physics, theoretical biology, and neurosciences. It is
remarkable that aspects of pattern discovery have only recently been explored in the field of genetics and bioinformatics. There is now a growing collection of investigations in bioinformatics attempting to investigate patterns, structures, and processes at every level of form, pattern, structure, function, interaction, and evolution through biological data objects.

The understanding of patterns, system biology, and network biology will be of crucial importance to the scientific understanding of living systems. However a full understanding of a living system requires further understanding the system’s pattern, structure, and process. A new synthesis of living systems was introduced by Capra [1]. The key idea of his synthesis is to express the key criteria of a living system in terms of three conceptual dimensions, pattern (autopoiesis), structure (dissipative structure), and process (cognition).

3.1. Autopoiesis-The Pattern of Life
Autopoiesis literally means "auto (self)-creation" and expresses a fundamental interaction between structure and function. The term was originally introduced by Humberto Maturana and Francisco Varela in 1973 [7, 8]. According to Maturana and Varela, a living system continuously produces itself. Autopoiesis is a network pattern in which the function of each component involves with the production or transformation of other components in the network. The simplest living system we know is the biological cell. The eukaryotic cell, for example, is made of various biochemical components such as nucleic acids and proteins, and is organized into bounded structures such as the cell nucleus, various organelles, a cell membrane and cytoskeleton. These structures, based on an external flow of molecules and energy, produce the components which, in turn, continue to maintain the organized bounded structure that gives rise to these components.

An autopoietic system is to be contrasted with an allopoietic system, such as a car factory, which uses raw materials (components) to generate a car (an organized structure) which is something other than itself (a factory). More generally, the term autopoiesis resembles the dynamics of a non-equilibrium system; that is, organized states that remain stable for long periods of time despite matter and energy continually flowing through them. From a very general point of view, the notion of autopoiesis is often associated with that of self-organization. However, an autopoietic system is autonomous and operationally closed, in the sense that every process within it directly helps in maintaining the whole. Autopoietic systems are structurally coupled with their medium in a dialect dynamics of changes that can be called sensory-motor coupling. This continuous dynamics is considered as knowledge and can be observed throughout life-forms [6, 9, 11]. Mathematical models of self-organizing networks were known as cellular automata - a powerful tool for simulating autopoiesis networks. A cellular automaton is a collection of "colored" cells on a grid of specified shape that evolves through a number of discrete time steps according to a set of rules based on the states of neighboring cells. The rules are then applied iteratively for as many time steps as desired. Von Neumann was one of the first people to consider such a model, and incorporated a cellular model into his "universal constructor." Cellular automata were studied in the early 1950's as a possible model for biological systems [13]. The simplest type of cellular automaton is a binary, nearest-neighbor, one-dimensional automaton. Such automata were called "elementary cellular automata" by S. Wolfram, who has extensively studied their amazing properties [13]. The theory of cellular automata is immensely rich, with simple rules and structures capable of producing a great variety of unexpected behaviors.

3.2. Dissipative Structure-the Structure of Living Systems
The term dissipative structure of a living system was coined by Ilya Prigogine who pioneered research in the field of
thermodynamics in [9]. A dissipative structure is a thermodynamically open system to the flow of energy and matter. A dissipative structure is operating far from thermodynamic equilibrium in an environment with which it exchanges energy and matter. Prigogine describes a living system as a dissipative structure and is both structurally open and organizationally close. Matter continually flows through it, but the system maintains a stable form, and it does so autonomously through self-organization. Simple examples of dissipative structure include convection, cyclones, and hurricanes. More complex examples include lasers, Bénard cells, the Belousov-Zhabotinsky reaction, and at the most sophisticated level, life itself. The vast network of metabolic processes keeps the system in a state far from equilibrium and gives rise to bifurcations through its inherent feedback loops.

3.3 Cognition—the Process of Life
The concept of cognition is closely related to such abstract concepts as mind, reasoning, perception, intelligence, learning, and many others that describe numerous capabilities of the human mind and expected properties of artificial or synthetic intelligence. Cognition or cognitive processes can be natural and artificial, conscious and not conscious; therefore, they are analyzed from different perspectives and in different contexts, in neurology, psychology, philosophy, and computer science. Cognition, according to Maturana and Varela, is the activity involved in the self-generation and self-perpetuation of living systems. In other words, cognition is the very process of life. In this new view, cognition involves the entire process of life—including perception, emotion, and behavior—and does not necessarily require a brain and a nervous system. At the human level, however, cognition includes language, conceptual thought, and all the other attributes of human consciousness. Mind is not a thing but a process—the process of cognition, which is identified with the process of life. The brain is a specific structure through which this process operates. The relationship between mind and brain, therefore, is one between process and structure. The brain, moreover, is by no means the only structure involved in the process of cognition. In the human organism, as in the organisms of all vertebrates, the immune system is increasingly being recognized as a network that is as complex and interconnected as the nervous system and serves equally important coordinating functions [1].

4. COGNITIVE COMPUTING
Cognitive informatics studies intelligent behavior and cognition. Cognition includes mental states and processes, such as thinking, reasoning, learning, perception, emotion, consciousness, remembering, language understanding and generation, etc. In the emerging theory of living systems mind is not a thing, but a process. It is cognition, the process of knowing, and it’s identified with the process of life itself. Cybernetics provided cognitive science with the first model of cognition—that is, as the manipulation of symbols based on a set of rules. The main themes of cognitive informatics encompass three categories of topics, i.e., cognitive computing, computational intelligence, and neural informatics outlined as the Theoretical Framework of Cognitive Informatics [12].

4.1 Evolving Cognition
Recently biologists and philosophers have been attracted by an evolutionary epistemology. It was argued that our cognitive abilities are the outcome of organic evolution, and that, conversely, evolution itself may be described as a cognition process. Furthermore, it is argued that the key to an adequate evolutionary epistemology lies in a system-theoretical approach to evolution which grows from, but goes beyond, Darwin's theory of natural selection. Although random mutation and natural selection are still acknowledged as important aspects of biological evolution, the central focus is shifting from evolution to co-evolution. This is an ongoing dance
that proceeds through a subtle interplay of competition and cooperation, creation and mutual adaptation. In other words, a proper understanding of human evolution is impossible without understanding the evolution of language, art, and culture. We must turn our attention to the mind-process of life.

4.2 Neural Networks
The concept of a neural network appears to have first been proposed by Alan Turing in his 1948 paper "Intelligent Machinery". Historically, computers evolved from the von Neumann architecture, which is based on sequential processing and execution of explicit instructions. On the other hand, the origins of neural networks are based on efforts to model information processing in biological systems, which may rely largely on parallel processing as well as implicit instructions based on recognition of patterns of 'sensory' input from external sources. In other words, at its very heart a neural network is a complex statistical processor. Artificial neural networks are made up of interconnecting artificial neurons. Artificial neural networks may either be used to gain an understanding of biological neural networks, or for solving artificial intelligence problems without necessarily creating a model of a real biological system. Biological neural networks are made up of real biological neurons that are connected or functionally-related in the peripheral nervous system or the central nervous system. A biological neuron may have as many as 10,000 different inputs, and may send its output to many other neurons. A single neuron may be connected to many other neurons and the total number of neurons and connections in a network may be extensive. In the field of neuroscience, they are often identified as groups of neurons that perform a specific physiological function in laboratory analysis. The cognitive modeling field involves the physical or mathematical modeling of the behavior of neural systems; ranging from the individual neural level such as modeling the spike response curves of neurons to a stimulus, through the neural cluster level such as modeling the release and effects of dopamine in the basal ganglia to the complete organism on behavioral modeling of the organism's response to stimuli. In more practical terms neural networks are non-linear statistical data modeling or decision making tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. Whilst a detailed description of neural systems is nebulous, progress is being charted towards a better understanding of basic biological mechanisms.

5. CHALLENGES AND PERSPECTIVES
From a scientific perspective discovering how the brain thinks is a major undertaking in the history of mankind. Bioinformatics provides computational and experimental tools to study the biological patterns, structures, and functions. Cognitive informatics investigates the internal information processing mechanisms and process of life-cognition. “Understanding the human mind in biological terms has emerged as the central challenge for science in the twenty first century. We want to understand the biological nature of perception, learning, memory, thought, consciousness, and the limits of free will,” as Kandel put it in [3] “Thus, we gain from the new science of mind not only insights into ourselves-how we perceive, learn, remember, feel, and act-but also a new perspective of ourselves in the context of biological evolution.” “The task of neural
science is to explain behavior in terms of the activities of the brain. How does the brain marshal its millions of individual nerve cells to produce behavior, and how are these cells influenced by the environment...? The last frontier of the biological sciences – their ultimate challenge – is to understand the biological basis of consciousness and the mental processes by which we perceive, act, learn, and remember.”

REFERENCES: