Edward R. Dougherty,
Gustavo A. Stolovitzky, Ioan Tabus,
and Xiaodong Wang

Signal Processing in Genomics

Most likely, many engineers do not associate basic signal processing topics such as state-space models, steady-state analysis, context-free grammars, control policies, and Markov chain, Monte Carlo methods with medicine. We hope this special issue on genomic signal processing will change this perception.

Genomics involves the study of large sets of genes and proteins, with the goal of understanding systems not simply components. It constitutes a fundamental domain of systems biology. Translational genomics refers to the translation of scientific genomic knowledge into medicine. It constitutes a major effort in systems medicine. The cell can be viewed as a complex network of nucleic acids, proteins, and other biomolecules whose interactions govern many fundamental cellular processes and functions. Genomics represents a fundamental shift in perspective, away from a focus on individual genes to the study of multivariate gene activity within the framework of cellular governance.

Genomic signal processing (GSP) has developed in the wake of this shift and has been defined as the analysis, processing, and use of genomic signals for gaining biological knowledge and the translation of that knowledge into systems-based applications. The aim of GSP is to integrate the theory and methods of signal processing with the global understanding of functional genomics with a special emphasis on genomic regulation. The two major practical goals of GSP relate directly to medicine: 1) to discover families of genes whose products (messenger RNA and protein) can be used to classify disease, thereby leading to molecular-based diagnosis and prognosis; and 2) to characterize genetic regulation, and its effects on cellular behaviour and function, thereby leading to a functional understanding of disease and the development of systems-based medical solutions.

[GENOMIC SIGNAL PROCESSING HAS BEEN DEFINED AS THE ANALYSIS, PROCESSING, AND USE OF GENOMIC SIGNALS FOR GAINING BIOLOGICAL KNOWLEDGE AND THE TRANSLATION OF THAT KNOWLEDGE INTO SYSTEMS-BASED APPLICATIONS.]

Given that systems biology concerns systems, it is inevitable that systems theory will come to dominate systems medicine; however, the existence of today's large body of knowledge concerning systems will not be easily applicable. There are many daunting issues, and they will not be overcome without a deep understanding of the biological and mathematical issues confronting us. Biological systems are nonlinear, display large degrees of redundancy, possess highly distributed regulation, and consist of extraordinarily large numbers of variables. This complexity is exacerbated by the difficulty of obtaining appropriate data, noisy measurements, small samples, and the tendency of biological systems to seek state stability when perturbed. New methods will be required. In particular, much work needs to be done in obtaining optimal information from small samples, approximation of complex systems, and robust filter design. These are challenges that should excite the signal processing community and stimulate it to take on another great challenge. This one means getting at the core of what it is to be alive and taking on the most deadly and intractable human diseases.

We hope that this special issue motivates young (and like some of us, the not so young) engineers to take the time to learn the requisite biology to enter the field of genomic signal processing. The opportunities to do important seminal research abound. Success will lead to mitigation of human suffering. This is a goal worthy of our best young minds.

We would like to acknowledge Prof. Ray Liu of the University of Maryland for encouraging us to organize this special issue. [SP]