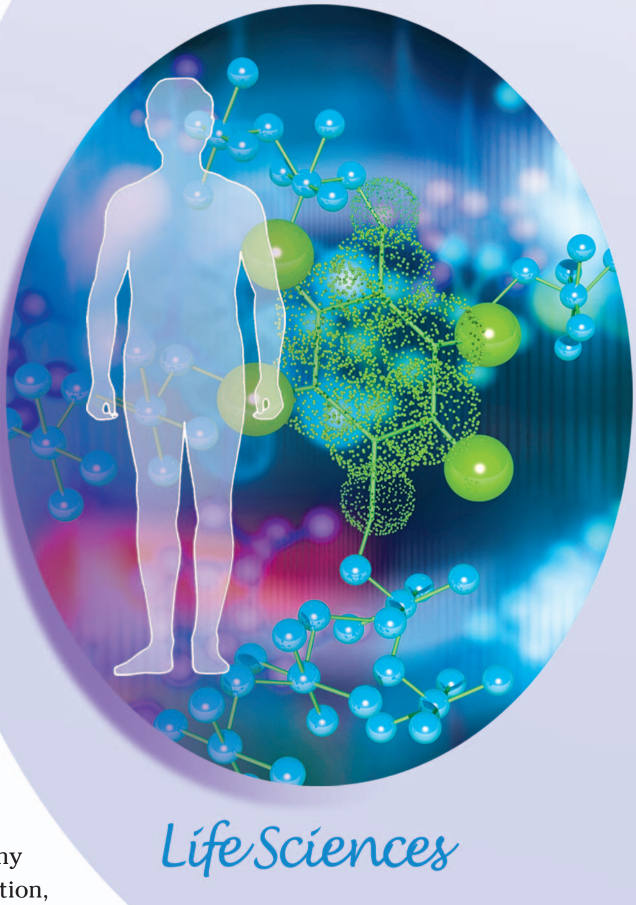
Opportunities in the Life Sciences

Mathukumalli Vidyasagar

Abstract

The purpose of this article is to review some recent developments in the life sciences area that are of interest to IEEE members, and to highlight IEEE's initiatives in the life sciences area.



It is widely believed that, just as the last half of the 20th century belonged to computing and communication, the first half of the 21st century will belong to biology. There is no doubt that a vast fraction of the technologies that have transformed life beyond recognition during the past fifty years or so have originated from the IEEE fraternity, more specifically the disciplines of computing and communications. Moore's law has now put more power into a simple desktop computer or laptop than was available in a 'work station' ten or fifteen years ago. Advances in data compression, data reconstruction, and signal processing have made it easy to achieve instantaneous worldwide communication from anywhere to anywhere. Smart phones can now combine features of a phone, an e-mail checker and a computing platform, all in a hand-held device. While there does not appear to be any end in sight to the computing and communications revolution, much of the excitement now surrounds the life sciences.

While it is true that science progresses through evolution and not revolution, often the imagination of the public at large is captured by a few watershed events. One such event was the sequencing of the human genome, which was simultaneously published by two groups in February 2001. In retrospect, the event was over-hyped. What was published was just a 'draft' human genome, which had up to 2% errors. To put that figure in perspective, it is widely accepted (though not yet fully verified), that the genomes of two humans have 99.9% overlap, while the genomes of humans and chimpanzees overlap 98%. Thus an accurate chimpanzee genome would still be a 'draft human genome with a 2% error.' Nevertheless, the popular press was full of futuristic prognostications about how 'personalized medicine' was just around the corner. The original

Digital Object Identifier 10.1109/MCAS.2012.2205973
Date of publication: 22 August 2012

Even if the cost of determining a genome were to follow below $1,000, the cost of *making sense* of the genome is still in the tens of thousands.

human genome sequencing project cost about $3.5 billion and took ten years. It was forecast that the cost of sequencing genomes would drop dramatically, with a biological version of Moore's law coming into effect. The idea was that one would determine the genomes of the entire planet, identify individual variations (called genotypes) and correlate those with various physiological classifications (called phenotypes). Indeed, the spectacular growth in the number of sequenced genomes bore out this optimism. Figure 1 shows the historical growth of the number of genomes deposited into genbank.[1] The latest available figures, as of April 15, 2012, show that there are 151,824,421 genomes encompassing a total of 139,266,481,398 base pairs. Due to the excitement surrounding the publication of the human genome, it is now common to refer to all biology after that event as 'post-genomic'.

In reality of course, the hype was at least ten years ahead of its time. Notwithstanding rosy projections, the cost of determining an individual genome is still around $5,000 to $10,000. The widely predicted $1,000 genome is not yet a reality. Companies whose business model consists of sequencing personal genomes (or building the devices that can sequence personal genomes) are in familiar territory for biotechnology start-ups, with huge burn rates and profitability not yet on the horizon.

Given that the total cost of health care for an American is around half a million dollars, it would appear to be worthwhile to spend a few thousand dollars per citizen, especially if this expenditure were to result in a significant savings in the total cost of providing health care for an individual over his/her lifetime. Even allowing for the fact that the cost of health care in the United States is roughly double that of the next highest country, and far higher than in almost all countries, it would still be worthwhile to construct genotypes of large sections of society, any society—*if* this raw data could somehow be translated into reliable conclusions.

Unfortunately this last step has remained elusive. While the technology for generating genomes is making good progress, the science for converting all this raw data into information, and information into knowledge, is simply not in place! Even for those diseases for which the underlying genetic foundational causes is well-understood, the cost of experts' time to apply this knowledge to raw genomic data is conservatively estimated at tens of thousands of dollars per patient. Hence, even if the raw cost of determining a genome were to drop below the $1,000 threshold, the cost of *making sense* of the genome is still in the tens of thousands, assuming that we as a community know what is to be done, which is a very dubious assumption.

It is not even clear whether genotype-phenotype correlation is indeed the right problem. The relatively small number of genome-wide association studies (GWAS) carried out thus far have not yielded anything significant. Indeed, in most cases the over-arching conclusion has been that 'environment' is at least as important as 'heredity' in determining one's health profile over a lifetime. This was true in the pre-genomic era, and apparently continues to be true in the post-genomic era as well. So the earlier optimistic dreams that, by studying a person's genome, it would be possible to predict a lifetime road map of potential health problems one or two decades ahead, and to prescribe appropriate health care (either preventive or curative), has proved to be quite unrealistic.

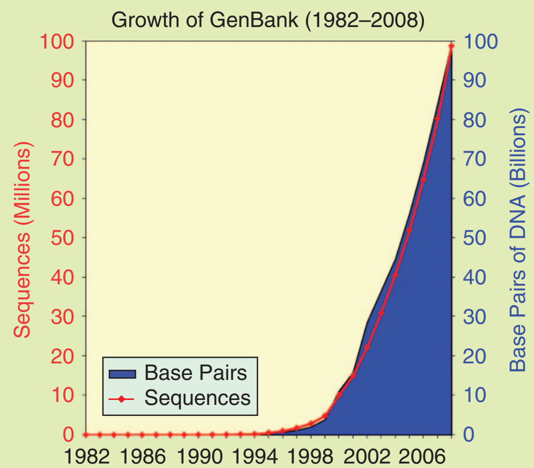


Figure 1. Historical growth of the number of genomes deposited into genbank. (Source: http://www.ncbi.nlm.nih.gov/genbank/genbankstats-2008/.)

[1]Source: http://www.ncbi.nlm.nih.gov/genbank/genbankstats-2008/

Mathukumalli Vidyasagar is the Cecil H. & Ida Green Chair in Systems Biology Science, The University of Texas at Dallas, 800 W. Campbell Road, Richardson, TX, 75080-3021 USA; e-mail: M.Vidyasagar@utdallas.edu.

The main challenge here is to turn the statement "This drug works for only 2–5% of the population" into "This drug works for only 2–5% of the population—but I can tell you which 2–5% that is!"

On the other hand, there does appear to be some light at the end of the tunnel in terms of *personal medicine*. While it is now not possible to predict who is at risk for what disease, especially many years ahead of time, there seems to be far greater scope for predicting what treatment regimen would evoke a good (or poor or even no) response in which individual, once a disease has set in. Perhaps the situation is most dramatically illustrated in cancer, in which no two manifestations are alike, even in the same organ. I like to paraphrase Leo Tolstoy and say "Normal cells are all alike. Every malignant cell is malignant in its own way."[2] Nowadays no researcher really talks about *the* cure for cancer. Rather, it is recognized that the landscape of cancer diagnosis and treatment will be like a patchwork quilt, in which the patient population is segmented into groups of perhaps as little as 2% of the patients, and each group is assigned a therapy that is most appropriate. The main challenge here is to turn the statement "This drug works for only 2–5% of the population" into "This drug works for only 2–5% of the population—*but I can tell you which 2–5% that is!*"

At this point one might be tempted to dismiss this entire topic as not being relevant to IEEE. But that would be a huge mistake. The challenge of personal medicine mentioned above is just one example whereby persons with a traditional engineering training and mind set can make quite fundamental contributions to the life sciences. Going beyond this specific area of application, it is clear that no discipline can survive in a static mode and electrical engineering (however defined) is no exception. From the standpoint of society at large, while computing and communication definitely have a large mind share, health care has a much bigger share. Health care costs account for roughly 17% of US GDP and climbing rapidly. The annual budget of NSF, at $6.859 billion in 2011, is dwarfed by the annual budget of NIH, at $31.987 billion in 2012. The two largest entities in the NIH, namely the National Cancer Institute (NCI) and the National Institute of Aging and Infectious Diseases (NIAID), have 2012 budgets of $5.196 billion and $4.916 billion respectively, comparable to that of NSF as a whole. Thus it is imperative for IEEE to reach out to the life sciences community.

[2] This paraphrases the opening sentence of *Anna Karenina*, which goes "Happy families are all alike. Every unhappy family is unhappy in its own way."

Due to the overarching importance of the subject, IEEE initiated a New Initiative in the Life Sciences in the calendar year 2010. I had the privilege of serving as the Chair of the committee to draw up a road map during that first year. From 2011 onwards, Prof. Bin He of the University of Minnesota and I have been serving as Co-Chairs of the committee. IEEE has historically had a presence in the life sciences area, defined broadly. Societies such as EMB (Engineering in Medicine and Biology) have been active in this domain for decades. It is noteworthy that as many as 29 Societies, Councils and Committees within IEEE are active in this domain. The following is the list of active entities.

- Antennas and Propagation
- Biometrics Council
- Circuits and Systems
- Communications
- Computational Intelligence
- Computer
- Control Systems
- Consumer Electronics
- Electron Devices
- Engineering in Medicine and Biology
- Geoscience and Remote Sensing
- Industry Applications
- Microwave Theory and Techniques
- Nanotechnology Council
- Nuclear Plasma Sciences
- Oceanic Engineering
- Photonics
- Power Electronics
- Reliability
- Robotics and Automation
- Sensors Council
- Signal Processing
- Solid State Circuits
- Society for Social Implications of Technology
- Systems Council
- Systems, Man and Cybernetics
- Ultrasonics, Ferroelectrics, and Frequency Control
- IEEE Committee on Earth Observations
- IEEE-USA, Medical Technology Policy Committee

There are as many as 30 IEEE Technical Committees dealing with life sciences-related matters, encompassing many different societies. There are eight publications dedicated to life sciences, and another fifteen publications that cover some aspects of life sciences. As

The translation of systems biology into a broader approach [will require] reversing the trend of the past 30 years, during which biology has become a discipline for people who want to do science without learning mathematics.

of the end of 2011, 41,414 articles published in IEEE are cited in Pubmed, the main database of the life sciences community.[3] In 2011, there were nearly 100 life sciences-related conferences under IEEE's auspices. Thus when the original Steering Committee was expanded to include representatives of all societies and other entities that had expressed an interest, the final team consisted of 35 members, divided into five sub-teams.

In 2012, the third and final year of the New Initiative in Life Sciences, a number of activities have been initiated, such a portal: http://lifesciences.ieee.org/ and a conference on grand challenges: http://lifesciences.ieee.org/grand-challenges-conference with many more activities being contemplated.

It is the fond hope of the Steering Committee that, once various activities have been nucleated, they would become self-sustaining. Ultimately it is individual initiative that will nurture such initiatives; a top-down fiat-driven growth model is not sustainable. Given that the *academic* life sciences community is much larger than the *academic* electrical engineering community,[4] the imperative for EEs to reach out to life scientists is greater than the other way around. Quite apart from this, my own past experience has been that it is relatively easy for engineers to pick up a *sufficient amount of biology* so as to interact effectively with the life sciences community, than the other way around.

In this connection I cannot resist quoting from the article by Dr. Marvin Cassman, former Director of Institute for Bioengineering, Biotechnology and Quantitative Biomedical Engineering, UC San Francisco, in *Nature*, 438, p. 1079, December 22, 2005:

> "Unfortunately, the translation of systems biology into a broader approach is complicated by the innumeracy of many biologists. Some modicum of mathematical training will be required, reversing the trend of the past 30 years, *during which biology has become a discipline for people who want to do science without learning mathematics*." (emphasis added)

I agree wholeheartedly with Dr. Cassman's observation. However, the lack of mathematical training in a typical biology or medicine program is not about to change anytime soon. Hence the EE community has to go more than 50% down the road in order to establish links with the life sciences community. I am 100% convinced that the returns would be well worth the effort.

Mathukumalli Vidyasagar was born in Guntur, India on September 29, 1947. He received the B.S., M.S. and Ph.D. degrees in electrical engineering from the University of Wisconsin in Madison, in 1965, 1967 and 1969, respectively.

For the next twenty years, he taught at Marquette University (1989-70), Concordia University (1970-80) and the University of Waterloo (1980-89). In 1989 he returned to India as the Director of the newly created Centre for Artificial Intelligence and Robotics (CAIR) in Bangalore, under the Ministry of Defence, Government of India.

In 2000 he moved to the Indian private sector as an Executive Vice President of India's largest software company, Tata Consultancy Services, where he created the Advanced Technology Center, an industrial R&D laboratory of around 80 engineers.

In 2009 he retired from TCS and joined the Erik Jonsson School of Engineering & Computer Science at the University of Texas at Dallas, as a Cecil & Ida Green Chair in Systems Biology Science.

In March 2010 he was named as the Founding Head of the newly created Bioengineering Department.

His current research interests are in the application of stochastic processes and stochastic modeling to problems in computational biology, control systems and quantitative finance.

He has received a number of awards in recognition of his research, including the IEEE Control Systems ("Field") Award in 2008 and Fellowship of the Royal Society, UK, in 2012.

[3] However, this is just a tiny fraction of the millions of papers in Pubmed.
[4] It would be a different comparison if persons in industry were to be included.