Crossroads
The ACM Student Magazine

Interview of Dr. Lee Hood

by ***Author Name***

Dr. Lee Hood founded the Institute of Systems Biology (ISB) in 2000. After graduate school at California Institute of Technology, he stayed with the university as a professor and pioneered the DNA sequencer. This invention serves as the foundation of molecular biology today and has made the Human Genome Project a reality. He has also played a role in founding biotech companies like Amgen, Applied Biosystems, and Rosetta.

Dr. Hood, how would you describe Systems Biology? Systems biology has 3 legs. The first leg, obviously, is the medical and biological aspects. The second leg is the instrumentation. I feel it is important to push state of the art in measurement tools because of the global nature of data. Finally, the third leg is comprised of computational tools. The goal of systems biology is to make simple models and be able to use them to predict how biological processes work.

Performing research in systems biology is an enormous challenge. Systems biology will be a new driver for biology in the twenty-first century, but I've noticed that whenever universities attempt to set up a systems biology department, they find that this new discipline is a huge undertaking and is often beyond their means. Each research team needs several researchers from different classical disciplines.

That's how we can be different at ISB. We're more than simply computers or bioinformatics. Systems biology enables us to validate the quality of data. It lets us model data either graphically or mathematically so that we can reduce the enormous data dimensionality and model diseases.

What kind of a relationship does ISB have with industry? We're starting to think of really big problems, so we need industry partners. These challenges are in the areas of predictive, preventive, and personalized medicine. To tackle such complexities we must integrate chemistry, biology, computer science, mathematics, physics, statistics, engineering, and other fields. Institutions such as ISB cannot support all the expertise needed, so we need to foster critical partnerships with industry.

This brings me to another important area, namely, funding. The federal funding situation will only deteriorate in the next few years, so we need to look for funding outside of government. We're not only interested in how to do science and in how to invent the tools, but also in how to organize science. We're interested in how to drive these big projects utilizing milestones, and we're interested in creating and pioneering new ways to fund these big projects.

How did you first get into the field of Systems Biology? I started as an assistant professor in 1970 at the California Institute of Technology (Caltech). Initially, I had two objectives: I wanted to study molecular immunology and immunobody diversification. In order to study these fields, I developed technologies and instrumentation to decipher the biological structures. I developed an automated protein sequencer, an automated protein synthesizer, and an automated sequencer and synthesizer for DNA. Then, at ISP, I developed ink jet technology to create micro arrays. This is the technology that Agilent still uses. While I was at Caltech, though, the university president told me that it was embarrassing to have all of this instrumentation in a biology department, so I moved all of the instrumentation to an independent company.

This had similar parallels in the NIH (National Institutes of Health). Initially, the NIH was one of the bitterest opponents to the genome project (the project that decoded human DNA). In order for the NIH to do this at all, they had to create a completely new entity there. As a consequence of the work I was doing, I really became convinced that the new challenges in biology were going to require the skills of researchers from multiple disciplines. So, around 1988 or 89, I approached the president of Caltech with the idea to set up a new department. The chemists and engineers supported it, but the biologists torpedoed it. I then went to University of Washington with Bill Gates' support.

At the University of Washington, I had the freedom that allowed me to begin to develop high throughput techniques that would be able to process more than a single protein at a time. I wrote my first grant proposal proposing systems biology in the late

80's, but I was severely criticized for it because no one understood it. Around 2000, it became evident to me that I wasn't going to be able to do this in a university setting, so I started ISB.

Was there a formative experience that opened your eyes to the need for the automated, high-throughput approach that you've championed? Bill Dryer at Caltech was my mentor. He proffered the sage advice to always operate at the leading edge because it's more interesting. He advocated the strategy that if one wants to invent a new area, one must invent new measurements. I realized that sequencers would be generating lots of data, and so I pushed the instrumentation that generated that data. I worked under the premise that if one could put all the genes of an organism on a chip, then we would be able to observe in a single inspection the changes as an organism or disease progresses.

Did anything surprise you along the way? If you want to understand biology, you need to understand the biological networks that mediate the processes. What surprised me was how great the computational and analytical challenges to manage this data were.

What do you see as the top challenges in informatics related to systems biology? I see a whole series of challenges. One is the question of how we can integrate the many bits and pieces of software programs that can analyze genes, proteins, and networks into an integrated system that biologists can use. There are an enormous variety of programs. How can we put them together to give global reach? A second challenge is the digital genome, and this is a core facet to the interdisciplinary goals. The computational questions evolving from this idea are fascinating. For example, can we take a digital genome and extract the biological networks while understanding their biological implications? We may not be able to predict how the environment and signaling modifies the digital genome; but, can we take all the genes and translate them in silico (in a computer simulation) into the 3 dimensional proteins?

This introduces a related problem, that of protein folding. Here, the problem is to see if we can take a linear representation of amino acids (the "primary structure") and fold them accurately in silico into a 3-dimension entity (the "tertiary structure").

Yet another top challenge is to investigate how to deal with the enormous amounts of data involved with in vivo visualization and then to mine this data for salient information. Everywhere you look, there are all kinds of fascinating challenges!

What risks are faced by systems biology? Scientists are terribly, terribly conservative! A lot of biologists are skeptical about the field of systems biology - just as they were with genome technology. We need to convince these skeptics. One way to do this is to determine how to create an environment where researchers can converse across disciplinary boundaries with colleagues that have the complementary skills and knowledge needed to make advances in this area. This is especially difficult in the current environment where sources of funding are drying up. In this kind of environment, institutions want to break up the large interdisciplinary centers.

There are also risks in ethico-socio-legal areas. An example of what I'm talking about here is the issue of genetic privacy. The United States is an environment where it is grotesquely complex to obtain permission to use samples. Because of these misguided bureaucratic efforts, I predict that most advances will not be made in the US. We see in Nature today that stem cell research has enabled paralyzed mice to walk again, but policy regarding this field is being guided by politics instead of by science, resulting in advances and research being blocked. So the risk is in finding ways of dealing with governments that not only are decreasing funding of medical research, but that are actively obstructing promising advances. Harvard's recent decision to go ahead and start a new line of stem cells anyway is great, but that then exposes them to the risk of not being able to get funding. Researchers in California now have the ability to get state funding, and making the lines available is a wonderful start, but won't solve the problem.

In my opinion the biggest risk of all is in K-12 education. We need to ensure that all citizens have enough knowledge to be able to make informed decisions on the issues. The United States is staggering in mediocrity. Of the top twenty developed countries we're nineteenth.

Systems biology requires the integration of many different types of biological data, for example, genomic, proteomic, and metabolomic. Do you see the semantic Web as being a key enabling technology for data integration? The semantic Web is a very important technology. I'm surprised it hasn't caught on quicker. I think that the biggest problem is trying to define what we're doing. Things in biology aren't very well defined. It's an exciting time. In a way, I wish that I were 20 years younger. On the other hand, I like being in this time and place. It's exciting to fund the younger scientists. I just don't want them to get discouraged; I want them to continue in this field because I think it is the most exciting career.