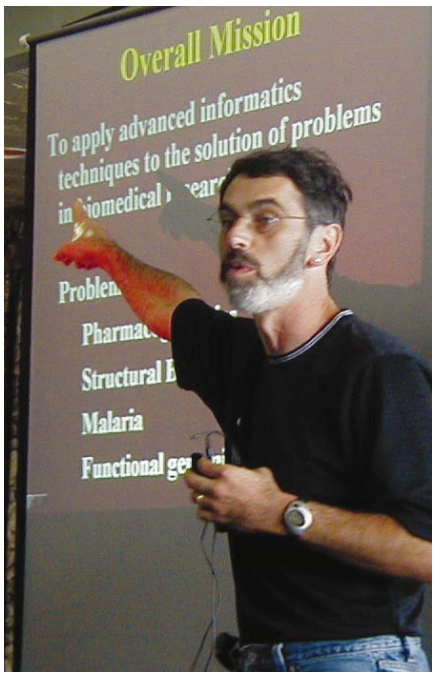
Training in a hybrid discipline

Courses to teach bioinformatics are starting to spring up all over North America. But the interdisciplinary nature of the subject means that there is a severe lack of experienced instructors, so the quality may vary between programmes, warns Potter Wickware.

A rarity five years ago, dedicated bioinformatics training programmes are now proliferating in North America. The International Society for Computational Biology lists more than 30 degree programmes in bioinformatics, covering 13 states and provinces in North America, and many more workshops and short courses. A total of 95 offerings worldwide for this marriage between biology and computer science are listed at Rockefeller University's website. Most of today's programmes tend to emphasize computer science but, with the emergence of functional genomics and proteomics, a shift in emphasis is now needed.



Russ Altman: computer science must recognize the challenges being issued by biology.

Francis Ouellette, director of the Bioinformatics Core Facility at the Centre for Molecular Medicine and Therapeutics in Vancouver, Canada, sees a need for greater flexibility in university rules if bioinformatics training is to improve and flourish.

"In Canada, and probably in the United States, undergraduate programmes in bioinformatics try to do a full core in biochem and a full core in computer science. But these overloaded programmes don't leave enough flexibility for the other courses that an undergraduate ought to be able to enjoy," Oullette says. "Even then, hybrid undergraduate bioinformatics degrees have incomplete biology and not enough computer science."

Universities should start bending their rules to make computer science more available to the biologists, Ouellette explains. "They should make it easy for biologists to get computer science courses, and vice versa, and even bring in other sciences as well," he says.

Another problem is the shortage of faculty with full bioinformatics credentials. "Many are biochemists who know a little computer science, or vice versa," Ouellette says. "It's a necessary compromise, and the hope is that there'll be enough cross-pollination for the programmes to work."

Why aren't there more senior bioinformatics investigators and instructors? As with so much in science, funding is key. "Only now are the funding agencies catching on that these projects are fundable and that it's worth attracting people to work in bioinformatics," Ouellette says.

GOING PAST HALFWAY

He believes that graduate degree programmes represent the only sustainable model for bioinformatics training. "I've been to biotech companies in Canada where their directors of bioinformatics turn out to be people who have taken my workshops, and I know how little they know," he says. "The workshops are getting people halfway there — but you're not going to get the innovators, the people who are thinking of new tools and new ways of thinking about the data, from a short course or workshop."

Russ Altman, director of the biomedical informatics training programme at Stanford University and president of the International Society for Computational Biology, thinks there's a critical distinction between programmes that train users and those that train tool-builders. "Tool users should take a short course, then use the tools on the job and continue getting ongoing training," he says. Tool builders, by contrast, need more training because they

Francis Ouellette sees graduate degree programmes as key to successful bioinformatics training.

must anticipate the problems that users will report when the tools begin to fall short.

Looking ahead, Altman sees deepening connections between bioinformatics, clinical medicine and databases, which will require an awareness of clinical issues by those who are being trained today.

"In basic science there will be more focus on complex systems, mathematical simulation and integration of data and knowledge," says Altman. "Structure, especially as a milestone on the way to functional prediction, will continue to grow in importance, as will interest in simulation."

As for the scope and quality of today's training programmes, Altman faults them on the same grounds as Ouellette. "Computer science, statistics and related disciplines are slow to recognize the challenges that biology is tossing out," he says. "Databases need to be integrated, new problems in statistics are piling up, and the computer engineers need to start looking at distributed systems in biology. The computer science and statistics departments need to start collaborating more with the biologists."

Tomorrow's practitioners will also need to be sophisticated about informed consent and patient privacy issues, given the increasing interconnectedness of databases and the need for large clinical trials.

The data arising from protein analysis will need skilled interpretation.

VARIAN

"The social–ethical side is lagging," Altman says. He is pressing for progress on this point as part of the Pharmacogenetics Knowledge Base project, a database project supported by the National Institutes of Health, and in the Stanford Informatics programme, where PhD candidates are focusing on ethical, legal and social issues in bioinformatics. But he also notes that passing new laws and regulations must proceed at a deliberate pace, while the technological advances that make them necessary are rushing forward unchecked. ■

Potter Wickware is a science writer in San Francisco.

NIH starts training

Although the US National Institutes of Health (NIH) has spawned many key bioinformatics tools, it has only recently launched a concerted effort to teach people how to use them.

This initiative has been made difficult because the types of data — and the tools to mine them — keep changing. In addition, top bioinformaticians are often tempted to leave academia for lucrative posts in industry.

Outside the NIH, foundations are being laid through a body of the institute that is only a few months old. The NIH's Center for Bioinformatics and Computational Biology (CBCB), established in May, has so far awarded grants to three US universities to set up pre-doctoral training programmes — the University of California, San Diego; Washington University, St Louis, Missouri; and Stanford University. Each is getting funds to establish up to three teaching posts, with the understanding that they try to bolster them with support staff.

Jim Cassatt, the CBCB's acting director, notes that recruiting for such positions is difficult because of competition with industry.

Inside the NIH, the National Center for Biotechnology Information (NCBI) plans to train groups at each of the NIH's 22 institutes in the finer points of the NCBI's bioinformatics software. Those people would, in turn, educate others working at the NIH.

The NCBI's David Wheeler, who is heading the training, says that most people who use bioinformatics software only scratch the surface of what it can do.

So far, five people have been trained, with four more in progress. If the programme is successful on the NIH campus, the NCBI may consider expanding it to help extramural researchers funded by the agency, but no such plans have yet been approved. Paul Smaglik

Web links

Canadian Bioinformatics Workshops ▸ **www.bioinformatics.ca**
Centre for Molecular Medicine and Therapeutics ▸ **www.cmmt.ubc.ca/ouellette**
Pathogenomics Project ▸ **www.pathogenomics.bc.ca**
Biomolecular Interaction Network Database ▸ **www.bind.ca**
Course list at Rockefeller University ▸ **linkage.rockefeller.edu/wli/bioinfocourse**
International Society for Computational Biology ▸ **www.iscb.org**