CHAPTER 9

Information Technology and the Biotech Revolution

Edward Skoudis

THE SECOND half of the 20th century is sometimes referred to as the information age, the space age, or the nuclear age. The first half of the 21st century may well be dominated by rapid advances in biological technologies. By leveraging the information technology base to control and manipulate biological systems, humans are poised for a biological revolution; the first fruits of such technologies are already available. Bioinformatics has yielded improved crop performance and a better understanding of some diseases for optimizing medical treatments, while brain-computer interfaces and neural prosthetics are helping disabled people interact with their surroundings more effectively.

The integration of information technology and biology is a long-time staple of science fiction dystopian visions that portray computers linked to biological systems that somehow supplant people or otherwise wreak havoc on society. While such technological integration could be a cause for concern, its benefits could also be quite profound in improving human lives and increasing economic efficiencies and progress. The purpose of this chapter is to survey advancements in the integration of information technology with biological systems to help the reader understand the evolutionary trends in this convergence of technologies.

Biology and cyberspace both deal with the interactions of a multitude of interconnected simpler elements, giving rise to many analogies between the two fields. Information technology has exhibited increasingly biological attributes as computers become more powerful, networks grow in size, and our overall cyberspace infrastructure becomes more complex. Some cyberspace terminology and analytical methods are lifted directly from biology. The term *virus*, first applied to self-replicating malicious software by researcher Fred Cohen in 1984, is now in the common vernacular.[1] Researchers refer to *outbreaks* and *inoculation*; experts caution against the dangers of a software *monoculture* without *diversity*, because of its lack of resilience to *digital pathogens*.[2] *Evolutionary algorithms* are used to find optimal answers to math problems by having software apply iterative, *evolutionary* steps to reach an answer.

However, the convergence of biology and information technology goes considerably beyond mere analogies of how complex systems behave. Increasingly, researchers are leveraging information technology to analyze, manipulate, and control biological systems. Research on many fronts includes two areas of particular interest that are the focus of this chapter: bioinformatics and computer-neuron interfaces.

Bioinformatics

Bioinformatics is a term sometimes used interchangeably with the phrase *computational biology* to refer to the application of modern statistical and computer science techniques to solving biological problems. Current bioinformatic analysis, modeling, and solutions may be focused on a molecular level to predict, for example, the interactions of deoxyribonucleic acid (DNA), proteins, and other vital biological chemicals. A DNA strand is made up of individual molecules, called nucleotides, which encode information. Different nucleotide sequences cause a cell to manufacture different proteins. Groups of these nucleotide sequences form genes. All of the genes for a given species make up that species' *genome*. One part of bioinformatics involves recording and analyzing nucleotide sequences and examining the implications—for the cell, the organism, and the whole species—of specific patterns in the DNA and the proteins it specifies. The bioinformatics field was born in the 1960s, when biologists began using mainframe computers to analyze data from protein biochemistry. With advances in computer technology and the rise of massively distributed computer systems, bioinformatic research has expanded significantly to model biological chemical interactions.

Bioinformatics Research Areas and Subfields

Within the broad definition of bioinformatics, research has branched out into many subfields; applications include gene finding and assembly, modeling of evolutionary processes, and analyzing mutations such as those related to cancer. The large amounts of data to parse and analyze when working with DNA, proteins, and related chemicals require massive computing capabilities. The genome of a species such as corn includes over a billion nucleotides, each storing a small amount of information about synthesizing and controlling the proteins that make up the corn. The gigabytes of data associated with this genome can be stored using an off-the-shelf computer. However, the data associated with cross- references within the genome could occupy terabytes of space. Beyond the mere storage of this genome and its cross-references, actual analysis of the data and simulation of its operation require very high-speed computing resources, usually networked computers calculating and searching in parallel.

To analyze a given species, the order of the nucleotides of its DNA must be recorded into data structures in a computer system, a process known as *sequencing*. Rather than sequencing DNA nucleotides one by one, a slow technique used in the past, many modern researchers use *shotgun sequencing*, a quicker technique that requires significant computational power from the bioinformatics arena. With this technique, the DNA of the organism under analysis is broken down at random into various chunks, which can be extracted in an efficient chemical process and entered into a distributed computer system. The computers then analyze the different piece-parts and reassemble them, sorting through millions of strands, removing overlapping segments, and finding and filling holes, to rebuild the original DNA structure. With significant computing power, the shotgun sequencing technique can improve speed over previous techniques by several orders of magnitude.

Sequencing is only the beginning, merely the capture of the raw data so it can be analyzed. Another area of bioinformatics research involves finding genes and predicting their expression in the final organism. For example, some cancer researchers are attempting to identify the gene or gene combination responsible for the expression of cancerous behavior in cells. Many of the techniques used in determining gene expression are noise-prone, as anomalous gene sequences that appear to match the desired results blur identification of correct answers. Researchers turn to the methods of bioinformatics, applying computer statistical analysis and filtering to cut through the clutter and find specific sequences associated with given behaviors or actions.

Bioinformatics studies also include the modeling of evolutionary traits and analysis of mutations by comparing genetic sequences between different and related species. For example, a *phylogenetic tree* is a depiction in a tree pattern showing similarities and deviations of natural gene mutation between different species to illustrate the relationships of different organisms.[3] This technique uses genetic similarities between species to create a sort of evolutionary map or family tree of relationships, which allows researchers to identify, at least tentatively, a genetic history of a set of species that are believed to have a common ancestor.

Finding and analyzing the common patterns in these genes to create the tree require significant computational power and have spun off their own subfield within bioinformatics called *computational phylogenetics*.

Some of the statistical techniques of bioinformatics are used in a sub-field called *comparative genomics*. Research in this area attempts to establish correspondence between genes of different species to help identify gene function in one species by comparing it to similar genes that have already been identified in another species. The Human Genome Project, for example, found correlations between human and mouse genes, which allowed researchers to identify gene properties in humans because the genes had the same purpose in both humans and mice. The processes involved with these techniques are heavily analytical and work with large data sets, using statistical techniques similar to those found in financial applications, including Markov chains, Monte Carlo simulations, and Bayesian analysis.

Detailed genetic analysis can allow a doctor to better understand a patient's health, analyze diseases in the patient, and predict how various drugs would affect that patient's biochemistry. With such information, the doctor can customize treatments and drugs to best match the patient's genetic profile, an approach sometimes referred to as personalized medicine. In such offerings, doctors identify biomarkers for a given disease and monitor progression of the disease on a molecular level. Treatments can then be based on individual data from one patient, rather than on generalized measurements from clinical trials over a large number of patients, the way most medical treatments are designed today. In addition, once a treatment has been administered, doctors can track a patient's response to the treatment, again on a molecular level. Personalized medicine techniques show significant promise for better diagnoses, more efficient drug development tailored to specific patients, and more effective targeted therapies. There are ethical and public policy implications associated with the rise of personalized medicine. To apply such techniques, some portion of the patient's genetic information must be gathered and scrutinized, perhaps before any disease is identified. This genetic database would most likely include raw gene sequences as well as protein information for the individual. Should an insurance or medical company have access to its customers' genetic sequences? Is it appropriate to calculate customers' insurance rates based on the possibilities of disease indicated by their genes? Some people might be effectively uninsurable, based on such information. In the next two decades, as researchers hone analytic capabilities and better understand the genetic composition or predisposition to medical maladies, such questions will become significant.

The bioinformatics areas described so far have focused on capturing and analyzing biological information using computers. Based on these results, however, researchers have gone further: they have altered the genes of a given species and moved genetic information from one species to another, giving rise to new properties in the resulting species. Such purposeful genetic manipulation is commonplace with commercial crops in many countries.[4] The United States alone accounts for approximately two-thirds of all genetically modified crops planted globally. Modified corn, cotton, and soybeans made up 40, 60, and over 80 percent, respectively, of each crop's acreage planted in the United States in 2007.[5] These crops may be modified to make them more resistant to drought, disease, and insects or to improve their nutritional content. While such genetically modified food is controversial—some European countries have placed severe restrictions on the import of such crops—it has increased the productiveness of cropland in countries where it is practiced.

Genetic modification is also possible with livestock, pets, and even humans. Gene therapy is an experimental technology used to alter an organism's genes, planting new information in them to treat a disease. Gene therapy was used successfully in 2002 on an infant with a defective gene that caused a complete shutdown of the child's immune system.[6] Doctors extracted unhealthy cells, added corrective genes to them, and then implanted the altered cells in the boy's body, where they were able to restart the immune system. Gene therapy has since been used in many similar cases and is now considered at least as effective as a bone marrow transplant in stimulating some patients' immune systems. Researchers are analyzing numerous other gene therapy techniques that could eradicate some forms of genetic disease and possibly even curtail some nongenetic diseases by increasing an organism's immune function.

Significant ethical dilemmas are raised with one generation's ability to alter the genetic makeup of all generations that follow. Who should decide which changes to make and which to forbid: government agencies, commercial interests, or parents? If we can make people who are smarter, healthier, and longer lived, should we? Can and should humans be altered to make them more docile or perhaps more aggressive? Will improved genetic traits be available only to those who can pay for them, or will they be shared across a society? Should genetic modification techniques be applied to military uses, either from a biological warfare perspective or to improve the characteristics of warfighters? Because DNA operates like a computer programmed via its nucleotide sequences, it is quite possible that given changes will work fine for many generations but will eventually cause problems, just as a computer program may run appropriately hundreds of times before crashing. Who, then, should be responsible for unanticipated negative consequences of genetic alterations?

Over the next several decades, societies around the world will have to contend with ethical dilemmas such as these that genetic manipulation poses. Such deliberations will not happen independently within each country: even if one country decides to boycott genetic manipulation while others endorse or encourage it, the boycotting country may fall behind its competitors technologically and economically. Moreover, since people, animals, and crops move across nation-state borders, a country that attempts to eschew genetic changes may not be able to keep such changes made elsewhere from crossing its borders and planting altered genes among its population. Altered genetic information already flows across borders, as pollen from genetically modified crops blows in the wind and pollinates unmodified crops hundreds of miles away, resulting in genetically modified descendants. Such issues and their importance will certainly increase.

Research Tools and Organizations Associated with Bioinformatics

At the root of bioinformatics lies the computer technology in which the biological information is stored and analyzed. The bioinformatics research community has developed and released many free, open-source, and commercial software utilities and online services for use by researchers around the world. Some of these tools are available from the National Center for Biotechnology Information (NCBI), operated by the U.S. National Library of Medicine and the National Institutes of Health. One of the most powerful and widely used databases, offered free by NCBI, is the Basic Local Alignment Search Tool, a service that can search for regions of similarity between biological sequences in various species, including humans, mice, rats, numerous plant species, and several microbes. With a Web front-end open to anyone with Internet access, this database allows for searches of nucleotides and proteins.[7] NCBI also offers the Entrez Protein Clusters database to help researchers find protein sequence and function similarities between species, and the Database of Genotype and Phenotype to help "elucidate the link between genes and disease." Such tools rely on data mining techniques to pull meaningful information from large masses of data.

For commercial purposes, searching through vast amounts of data to find relevant results has been a major focus of Internet search engines. The skills and technologies honed to locate Web pages on the Internet are now being repurposed to the bioinformatics field. Google, in particular, has taken a keen interest in bioinformatics research and the use of Google technology to improve search capabilities. Teaming with Craig Venter from the Human Genome Project, Google has set out to apply its search algorithms to genetic data and to create an entire database of genetic information available for all to access. Google founders Larry Page and Sergey Brin have recognized the relationship between search engine technology and data mining of genetic information, spearheading genetic storage and search projects that could portend genetic information stored in a Google database and searchable with Google technology.[8]

A commercial company focused on analyzing genetic information, called 23andMe, was started in 2006. It seeks to offer people insight into their own ancestry, genealogy, and inherited traits based on the information in their genes, stored on the 23 paired chromosomes in each human cell.[9] 23andMe aims to help "put your genome into the larger context of human commonality and diversity," based on genetic information. Customers provide a sample of saliva, which is sequenced and analyzed for a fee. Then they are given detailed reports on their genetic predisposition to various traits, along with commentary about how common each trait is. Reports also include explanations for tendencies in customers, such as preferences for certain foods or character traits that are associated with specific genetic sequences. Aiming to provide more value to customers, 23andMe offers recommendations for changes in lifestyle that could lower the chance of contracting a genetically predispositioned ailment, such as heart disease. Numerous other small startups are being created to offer similar services derived from the growing bioinformatics industry.

Connecting Neurons to Computers: Neuralprosthetics and Brain-Computer Interface

A second major area of convergence of biology and information technology involves the interface between biological neural systems and computers. Such interconnections fall into two general categories: neuralprosthetics and brain- computer interfaces (BCIs). Each could eventually allow humans to control machines using thought or improve mental faculties with computer augmentation. Neuralprosthetic technologies focus on tying computer equipment into nerves outside of the brain, such as those in the ear, eye, spinal cord, arm, or leg. BCIs involve direct interface between computers and the motor, auditory, or visual cortexes of the brain itself.

Today, BCIs are highly experimental, but some neuralprosthetic technologies are already in general use as medical devices. Cochlear implants, for example, apply computer technology to help translate sound into neural signals to allow deaf people to hear. They represent one of the greatest areas of success and commercialization of neuralprosthetics; over 100,000 patients were estimated to use these devices worldwide in 2006.[10] Artificial intelligence researcher Ray Kurzweil suggests that additional processing and computing technology could augment these implants in the future, enabling them to translate languages in real time or whisper the definitions of unfamiliar terms to the implant user.[11]

A promising near-term application of BCI is the “brain pacemaker,” which injects electrical signals to alleviate symptoms of Parkinson’s disease or clinical depression in patients who do not respond to traditional medications. Experimental studies since 1998 have shown improvement in patients attributable to experiments with the technique.[12] Early experiments implanted such devices directly into patients’ brains, representing a medical application of BCI. Other researchers have postulated that such effects might be achievable by connecting instead to spinal cords or cranial nerves, putting the devices into the category of neuralprosthetics.

Most of the experiments to date with BCI and neuralprosthetic devices have focused on one-way communication between the computer and neurons, either writing data into neurons, such as with a cochlear implant or brain pacemaker, or reading data from neurons, such as with experimental brain control of robotic arms or computer cursors. While one-way interfaces are far simpler to construct, they offer no direct feedback, and the control they offer is much more limited than with two-way devices, which would be far more complex but will likely develop as an important area of BCI and neuralprosthetics research.

Researchers have experimented with a variety of techniques for deploying both BCIs and neuralprosthetics. The most direct route involves implanting probes directly into the gray matter of the brain, with electrodes protruding from the skull. Researcher Miguel Nicolelis at Duke University began employing this technique in 2000 in experiments that have allowed monkeys to control computer cursors and/or robotic limbs.[13] Another approach involves brain implants, placed inside the skull and touching the surface of the brain but outside of the brain’s gray matter, that measure electrical signals. Sometimes called electrocorticography (ECoG), this technique was used in 2006 to enable a teenage boy to play a video game by merely thinking about controlling the movements of a joystick.[14] To obviate the need for the “tether” associated with gray matter interfaces or with ECoG inside the skull, researchers are working on plans to implant a wireless transceiver in the brain that would transmit signals through the skull. This would allow patients to be more mobile, but it imposes severe size constraints on the local sensor equipment.

For each of these approaches, several experimental technologies have been devised for connecting individual neurons to computer equipment. One method is to graft the sensor directly into the neuron itself, where it would sense the signals transmitted by that neuron. In a less intrusive approach, the sensor would touch the surface of the neuron to detect its electrical state. Some researchers are analyzing use of lasers to bounce light patterns off of individual neurons to measure the changes in their reflectance that occurs as the neurons fire. This technique, called light reactive imaging, might have a less destructive effect on the neurons than other methods.

Other approaches to BCI and neural prosthetics would avoid the cost and invasiveness of inserting sensors into the body, because surgery would not be required. Some researchers are working on probes that would make contact with skin in an approach called electroencephalography (EEG); they would require no surgery but might involve shaving certain areas of the scalp. Because the skull dampens electrical signals from inside the brain, significant amplification and noise reduction would be required to get meaningful information. This technique cannot pinpoint specific neural activities but instead gathers information from a large number of neurons. While this method presents challenges, it is noninvasive and relatively low cost. An approach that is essentially “EEG at a distance” would involve sensors that can monitor brain activity remotely. This technique would require no physical contact with the patient, but it would constrain mobility of the patient so signals could be measured across a room, and it would be limited by noise from the outside environment.

Another promising approach that is neither a BCI nor a neuralprosthetic device pulls information from muscle movements, not from neurons. It employs electromyographs (EMGs), sensors connected to or embedded in muscle tissue. By measuring small changes in muscle tension, such technologies have allowed severely paralyzed patients to interact with computer equipment so they can control a voice synthesizer, computer keyboard, or other technology by moving muscles that are not paralyzed, such as their eyelids or eyes.

Organizations Associated with BCI and Neuralprosthetics

Numerous organizations are involved in BCI and neuralprosthetics work, including Duke University, Brown University, Stanford University, and the University of Sussex in England. A prominent BCI company is Cyberkinetics Neurotechnology Systems, which is experimenting with devices that let humans control computer cursor movements using sensor arrays implanted in the motor cortex of the brain. The device, about the size of an aspirin, provides approximately 100 electrodes to interface with the brain. In experiments in 2005, a quadriplegic patient implanted with the technology controlled an artificial hand, moved a computer cursor, turned lights on, and changed television channels by simply thinking about these actions. The long-term goal is to develop devices such as “thought-controlled” wheelchairs and prosthetic limbs.

Neural Signals, another major BCI and neuralprosthetics company, focuses on restoring speech by means of both invasive and noninvasive technologies. Researchers are experimenting with probes into the speech center of the brain, called the Broca’s motor area. Based on the activity they detect, a computer generates 1 of the 39 phonemes that constitute the fundamental sounds of spoken English. With training, patients may be able to make the machine speak for them simply by thinking. A noninvasive approach for discerning a patient’s thoughts and converting them into speech or action involves EEG signals from probes taped onto a patient’s scalp. Already, such devices can differentiate the thoughts “yes” and “no” within 10 seconds.[15] The company has also commercialized EMG sensors that, when placed on the skin of speech- and motion-impaired patients, can discern small motor movements of the eye and other muscles to control computers.

With its focus on helping disabled people, BCI and neuralprosthetics research has generated far less controversy than genetic manipulation. While genetic manipulation could affect all subsequent generations of life, and altered genes might spread unchecked, BCI and neuralprosthetics are more controllable technologies; skilled personnel apply them deliberately. Yet if these technologies continue to advance and are widely deployed, they too could result in significant changes to human civilization. Some researchers posit a future with thought-controlled robotic limbs and augmentation of human senses with computer devices. Humans with such gear embedded into their bodies could have superhuman senses or superhuman strength and endurance. Futurists and science fiction writers have also envisioned computing devices that would augment the brain to offer, for example, vast increases in human memory. A photographic memory might become the norm among people who can afford the technology, profoundly increasing economic disparities in society, giving rise to an elite class with improved capabilities baked into their genes or built into their skulls. Humans with such implants might be able to look up facts on the Internet merely by thinking about them or conduct conference calls with other people without any outside equipment, all inside their heads. Military fighters with superhuman strength or intelligence delivered by BCI and neuralprosthetics technologies could dominate in combat, although opposing forces could find a way to neutralize or counter the technology with, for example, computer network attack and exploitation techniques, just as other "ultimate weapons" have historically been countered by lower tech asymmetric strategies.

Conclusion

The applications of the integration of biology and information technology are just the start of even more profound capabilities. Taken to their extreme, biology and information technology used together could transform what it means to be human. In a "trans-human" future, people might develop a new species or groups of species to succeed humanity. However, these transitions are likely to be gradual and taken for granted by most people, as their lives improve with the accumulation of technological changes. Drastic setbacks are certainly possible: crops might fail due to poorly planned genetic manipulation, or machine-augmented warfighters might cause significant damage. Nonetheless, the advantages offered by biological technology are likely to make such advances inevitable. Noted physicist and technology thinker Freeman Dyson, referring to technologies associated with genetic manipulation as "green technologies," says:

> Before genetically modified termites and trees can be allowed to help solve our economic and environmental problems, great arguments will rage over the possible damage they may do. . . . I am not saying that the political acceptance of green technology will be quick or easy. I say only that green technology has enormous promise for preserving the balance of nature on this planet as well as for relieving human misery.[16]

In the early 21st century, information technology and cyberspace provide an ideal base for a technological revolution in biology. This next revolution could have major impact on the way we live our lives and what our lives really are.

---

[1] Fred Cohen, "Experiments with Computer Viruses," 1984, available at <http://all. net/books/virus/part5.html>; Judith Klein-Seetharaman, "The Use of Analogies for Interdisciplinary Research in the Convergence of Nano-, Bio-, and Information Technology," National Science Foundation Report on Societal Implications of Nanoscience and Nanotechnology, 2004, available at <www.cs.cmu.edu/~judithks/ Klein-Seetharaman.2005.pdf>.

[2] Daniel Geer et al., "CyberInsecurity: The Cost of Monopoly," September 2003, available at <http://cryptome.org/cyberinsecurity.htm>.

[3] "Phylogenetic tree," *Wikipedia*, available at <http://en.wikipedia.org/wiki/Phylogenetic_tree>.

[4] Argentina, Brazil, Canada, China, South Africa, and the United States accounted for 99 percent of genetically manipulated crops in 2003; Pew Initiative on Food and Biotechnology fact sheet, "Genetically Modified Crops in the United States," August 2004, available at <www.pewtrusts.org/news_room_detail.aspx?id=17950>.

[5] U.S. Department of Agriculture Economic Research Service Data Sets, "Adoption of Genetically Engineered Crops in the U.S.," July 2007, available at <www.ers.usda.gov/ Data/BiotechCrops>.

[6] Judy Siegel-Itzkovich, "Scientists Use Gene Therapy to Cure Immune Deficient Child," *British Medical Journal* 325, no. 7354 (July 6, 2002), available at <www.pubmedcentral.nih. gov/articlerender.fcgi?artid=1123542>.

[7] See Basic Local Assignment Search Tool at <www.ncbi.nlm.nih.gov/BLAST>. Harvard University provides a similar tool called QueryGene, available at <http://llama.med. harvard.edu/~jklekota/QueryGene.html>.

[8] David A. Vise and Mark Malseed, *The Google Story* (New York: Bantam Dell, 2005), chapter 26, "Googling Your Genes."

[9] 23andMe Web site, available at <www.23andme.com>.

[10] University of Michigan news release, "New Cochlear Implant Could Improve Hearing," February 6, 2006, available at <www.umich.edu/news/index.html?Releases/2006/ Feb06/r020606a>.

---

[11] Ray Kurzweil, *The Age of Spiritual Machines* (New York: Penguin Books, 1999).

[12] Kristin Weidenbach, "Pacemaker for the Brain May Offer Control for Severe Depression," *Stanford Report*, October 11, 2000, available at <http://news-service.stanford.edu/ news/2000/october11/brain_pace-1011.html>; BBC Report, "Health: Latest News 'Pacemaker' for Parkinson's Disease," May 20, 1998, available at <http://news.bbc. co.uk/2/hi/health/97057.stm>.

[13] Elizabeth A. Thomson, "Monkey Controls Robotic Arm Using Brain Signals Sent over the Internet," MIT News Office, December 6, 2000, available at <http://web. mit.edu/newsoffice/2000/monkeys-1206.html>; Duncan Graham-Rowe, "Monkey's Brain Signals Control 'Third Arm'," *New Scientist*, October 2003, available at <www. newscientist.com/article/dn4262-monkeys-brain-signals-control-third-arm.html>.

[14] Tony Fitzpatrick, "Teenager Moves Video Icons Just by Imagination," Washington University in St. Louis News and Information, October 9, 2006, available at <http:// news-info.wustl.edu/news/page/normal/7800.html>.

[15] Neural Signals Web site, available at <www.neuralsignals.com>.

[16] Freeman Dyson, "Our Biotech Future," *The New York Review of Books*, July 19, 2007.