BLURRING THE BOUNDARIES

Richard Williams MBCS CITP **discusses the area of computational biology, a discipline where computer science meets biology.**

The idea of modelling a biological process computationally is almost as old as the electronic computer itself. Alan Turing, in 1952, modelled the movement of chemical substances in a cell, along with a number of basic chemical reactions associated with these chemicals, and their movement between neighbouring cells.

More recently, the separate disciplines of computer science and biology have begun to formally merge through the interdisciplinary field of computational biology. This new and growing discipline is vast, encompassing all fields of biology that have synergistically joined forces with computer science.

The main objective is to gain a better understanding of the underlying processes and mechanisms of life. During the latter half of the 20th century, biology was dominated with experiments that looked at increasingly smaller and smaller components (the reductionist approach). This has been very successful and has generated a large amount of information about individual cellular components and their functions. In fact, the last two decades have seen an acceleration in the reductionist approaches through the emergence of high-throughput technologies.

Genetic identity

The landmark Human Genome Project (HGP) was completed in 2003 and is believed to have been a success as it mapped out the genetic identity of our species. The vast quantities of data generated through this project have uncovered a crucial problem, however, the need to deal with the complexity of the studied systems, requires development of technologies and tools to synthesise knowledge from the data.

Thanks to advances in computing power, information technology, and a large number of new biological techniques for gathering data, it is believed that biological sciences are heading for a revolution. A major concern following completion of the HGP relates to the concept attributed to Aristotle that 'the whole is more than the sum of its parts'.

This suggests that dissecting an organism, tissue or cell into increasingly smaller subunits (such as DNA sequences) and hoping to be able to piece it back together afterwards will not work if the

doi:10.1093/itnow/bwr038 ©2011 The British Computer Society

underlying mechanisms and relationships between these subunits are anything but linear and static.

Organise data

Real knowledge is much more than just data, as it arises from not only our ability to take measurements, but also from the intellectual frameworks that we create to organise and reason about data.

Modelling is one of the most powerful and widely used forms of intellectual framework in science and engineering, and advances in computer hardware and programming languages (and techniques) have facilitated a new and rapidly growing area of computational biology that focuses on modelling and simulation of complex biological systems. Computational biology has two distinct branches: knowledge discovery or data-mining, which used to be termed Bioinformatics, and tries to generate hypotheses by extracting the hidden patterns from huge quantities of experimental data and simulation-based analysis.

Computer modelling and simulation allows the investigation and analysis of aspects of complex biological systems that we are unable to observe or understand directly.

This may be due to difficulties in observing system dynamics due to extremely short or indeed large time scales, to the magnitude of the system to be observed (e.g. whole populations), or to the location or complexity of the system.

Simulation attempts to predict the dynamics of systems so that the validity of the underlying assumptions behind the models can be tested. Simulation can also be used to generate predictions for further biology laboratory-based (so called wet-lab) experiments, and to explore questions that are not yet amenable to wet-lab experimental inquiry.

This approach of performing experimentation in a computer has been termed in silico experimentation. Models and simulations are generally based on the tools available at the time and early examples relied heavily on cellular automata and differential equations. With the advances over time to computing power, hardware (particularly graphics), and the widespread use of object-oriented programming languages, biological models evolved to utilise parallel processing through advanced cellular automata and more recently agent-based modelling and simulation.

An example of the benefits such an approach offers is the potential to accelerate drug discovery relating to therapeutic treatment of diseases.

Just as the biological sciences have their own sub-disciplines, so to with computational biology. A good example of a sub-discipline of biology is immunology, which relates to the function of the immune system.

Parallels are made in the computing world with the new and rapidly growing field of computational immunology, which uses modelling and simulation to understand how the immune system functions and how diseases spread when the immune system does not function correctly.

Infections or disease

Computational immunology has the potential to facilitate more accurate prediction of the properties and behaviour of living organisms when experiencing infection or disease. The development of computational models of diseases and their response to computer generated perturbations allows the initial testing of hypotheses to be performed in silico, which may then be validated through subsequent wet-lab experimentation.

An example of the benefits such an approach offers is the potential to accelerate drug discovery relating to therapeutic treatment of diseases by reducing the number of potential hypotheses of drug targets through in silico experimentation. This approach yields only those hypotheses with the greatest probability of success during laboratory-based research and development and resultant clinical trials.

The immune modelling group in which I am researching for a PhD, uses agent-based modelling and simulation techniques to facilitate further understanding of a number of complex biological systems. These include diseases such as multiple sclerosis, through simulation of the animal equivalent experimental autoimmune encephalomyelitis; dynamics of white blood cells to form granuloma (spherical mass of cells to box in foreign substances); and my particular area of research [1] involving the NF-kB cell signalling pathway, which when dysregulated is involved in a whole host of diseases from arthritis, to cancer, to heart disease.

Other groups have made advances in modelling human organs. A 3D multi-scale agent based model of a heart has been developed that links the biochemical dynamics of individual cells up to the level of the whole organ in a quantitative and predictive manner [2]. This model enables reconstruction of the spread of excitation through the heart and makes it possible to reconstruct a number of life-threatening heart conditions such as arythmias within a computer simulation.

Modelling

Other organs such as the lungs, brain and blood system have been modelled and it is conceivable that these could one day be connected in the hope of developing a model of an entire human. In fact, a workshop held in Tokyo in 2008 laid down a grand challenge to create a virtual human in the next 30 years. This is only the beginning however. In the future we will begin to develop in silico organisms that are computer representations of their real world counterparts.

Due to the intrinsic complexity of biological systems, wet-lab experimental approaches alone will not uncover the design principles behind the functionality. A combination of laboratory-based and computational approaches is expected to resolve this problem.

The latter half of the 20th century has tentatively been termed by some scientists to have been the biological revolution. It is conceivable that the early 21st century may be deemed by future generations to have been the start of the computational biology revolution.

References

[1] http://www-users.cs.york.ac.uk/~rawillia

[2] Noble, D., 2002, Modelling the Heart: From Genes to Cell to the Whole Organ, Science, volume 295, pp.1678-1682

www.bcs.org