Computational Challenges of Systems Biology

Anthony Finkelstein, James Hetherington, Linzhong Li, Ofer Margoninski, Peter Saffrey, Rob Seymour & Anne Warner

CoMPLEX, University College London, Gower Street, London WC1E 6BT, UK

{a.finkelstein|j.hetherington|l.li|o.margoninski|p.saffrey|r.seymour|a.warner}@ucl.ac.uk

Systems Biology is an important and highly demanding area of interdisciplinary science. We argue that, as a discipline, computer science has much to contribute to meeting these demands. We outline what is meant by systems biology, and contrast it with bioinformatics. We then briefly review the computing contributions to the state-of-the-art. The paper uses a simple information model to outline the challenges for computer science in this new and largely uncharted area. An example is presented which illustrates the complexities that must be faced. We discuss modelling strategy and conclude by marking some staging posts for future progress.

Motivation

New applications of computing rarely attract much attention from computer scientists unless they pose novel computational challenges, stretch the state-of-the-art or open an unanticipated use of computing concepts. Bioinformatics is an example of an application that has attracted such attention. Bioinformatics is, in essence, the computing response to the molecular revolution in biology.

This revolution has reshaped the life sciences. We now understand the DNA sequence of many genes, up to whole genomes. We also understand the mechanics of much of RNA synthesis in exquisite detail. We know the genetic code for specifying amino acids so that the backbone of a protein can be directly predicted from DNA sequence information. Some of the complexities of RNA splicing, the means by which one gene can generate many RNAs and therefore proteins, are understood. We understand that DNA sequences, called promoters, determine which genes are expressed and that DNA binding proteins, called transcription factors, modify gene expression. We have a limited understanding of these transcription factors and the manner in which they can promote or inhibit the expression of particular RNAs. Knocking-out and over-expressing genes and RNAs have revealed how particular genes contribute to certain biological processes; it has also revealed substantial functional redundancy.

In the process of achieving this revolution in understanding we have accumulated very large amounts of data. The scale of the data, its structure and the nature of the analytic

task have merited serious attention from computer scientists and have prompted work in intelligent systems, data-mining, visualisation and more. It has also demanded serious efforts in large-scale data curation and worldwide infrastructure to support this. These approaches are collected together under the umbrella of Bioinformatics, which is perhaps the handmaiden of molecular biology. In tackling such problems computer scientists have the additional satisfaction of contributing to a scientific grand challenge.

Bioinformatics is only the first step in reshaping the life sciences. For further progress, we must return to the study of whole biological systems: the heart, the cardiovascular system, the brain, the liver — systems biology. To succeed we must combine information from the many rich areas of biological information. Alongside the *genome*, our knowledge about genes, we place the *proteome*, *metabolome*, and *physiome,* our information about proteins, metabolic processes, and physiology. We will bring these together to build an integrated physiology of whole systems. Systems biology is at least as demanding as, and perhaps more demanding than, the genomic challenge that has fired international science and excited the attention of the public. To achieve it will involve computer scientists working in close partnership with life scientists and mathematicians. By contrast with the molecular biology revolution, computer science will be proactively engaged in shaping the endeavour rather than clearing up afterwards!

The prize to be attained is immense. From 'in-silico' drug design and drug testing, to individualised medicine that will take into account physiology and genetic profile, systems biology has the potential to have a profound impact on healthcare and beyond.

The Role of Modelling

Even if we had a catalogue of all the gene sequences, how they are translated to make proteins, which protein can interact with which, and the way in which the protein backbones fold (into sheets, helices etc. with differing properties), we would not be able to put them into a functionally meaningful framework simply from the data. There are important reasons for this. All proteins are 'post-translationally' modified. Side chains are added (like sugars) to make, for example glyco-proteins, important constituents of cell membranes. These additions influence the shape and properties of proteins and hence their function and behaviour. Further, just because two proteins can, in principle, interact, it does not mean that they do so in real cells. Also, many functionally important, small molecules are synthesized by metabolism. For example, many neuro-transmitters are made by cells and are not translated from RNAs. Biological systems are so enormously complicated that however much we learn about them, a full simulation based on complete understanding will never be possible.

Thus, a bottom-up, 'data-driven' strategy, will not work — we cannot build an understanding of biological systems from an understanding of the components alone. What other approaches might be tried? At the heart of the emerging field of systems biology is modelling. We can use experimental information to build models at different biological scales, integrating these models to create an 'orchestrated' assemblage of models ranging from gross models of physiological function through to detailed models that build directly on molecular data. In principle these models should span from DNA and gene expression, through intracellular networks, on to cell-cell and trans-membrane signals to an organ level [Figure 1]. Most tenuously we might eventually construct such models at the organism level.

We thus introduce two key concepts for systems biology, methodologies that are forced upon us by the peculiar complexity of biological systems. First, the importance of simplification – biological complexity requires us to model, not simulate. Second, the importance both of modularity and of the integration of the modules – biological complexity requires us to break our systems down into manageable components, but also requires us to reassemble them, since behaviours can emerge which cannot be understood from the components alone. We will discuss these concepts with examples below.

The resulting models can serve to provide coarse grain prediction, can be used as a scaffold for our emerging understanding of the data, can identify gaps in our biological knowledge and, if the models are good, predict new behaviours that can then be explored experimentally. Iteration between model and experiment is the key to ensuring that models are realistic. This too can be difficult, since physiological studies of whole systems, where the individual components are monitored simultaneously, are extremely difficult, if not impossible, because a different technique can be required to study each component.

This agenda however poses some very serious challenges in the construction, integration and management of the models. Computer scientists are well placed to meet these challenges. Before we review them, we briefly summarise progress to date.

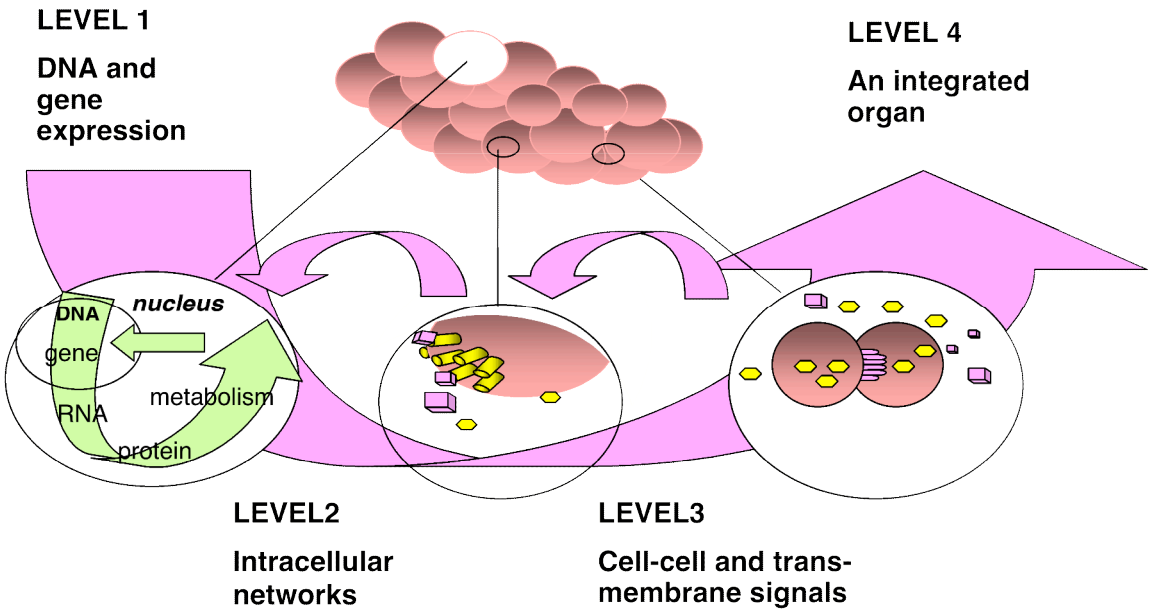

Figure 1: Building Models in Systems Biology

Modelling State-of-the-Art

The paradigmatic example of systems biology is the model of the heart developed by Denis Noble, Peter Hunter and colleagues [1]. This provides a computational model of the electrical and mechanical activity of the heart in health and disease. The model has been linked to sophisticated visualisations, particularly solid geometry models. The model has proven invaluable in developing an understanding of cardiac arrhythmia with consequences both for drug design and testing [2].

The model itself has grown from relatively simple beginnings in 1962 as an adaptation of the classic Hodgkin-Huxley squid axon model (one of the landmark achievements of modern biology) [3], to its current state involving hundreds of equations and adjunct models, such as a finite element model. Despite this sophistication and the large amount of effort it has consumed, the model only covers a small part of the mechanical, electro-physiological, and chemical phenomena of the heart. The heart model thus reveals not just what can be achieved but suggests the scale of the challenge that Systems Biology presents. The model has been the seed for the Physiome project [4] that collects and catalogues biological models and supports access to these models. It also provides web accessible databases of biological data that can potentially be linked to models.

There exist a plethora of 'stand-alone' models of various biological phenomena produced by different researchers. They are mostly relatively simple, although some are more sophisticated. One such is the bacterial model of Dennis Bray and colleagues [5]

that models flagellar motion (flagella are thin projections from cells) and chemo-sensitivity. Many models are provisional, in that they embed contested hypotheses about biological function or structure, or are otherwise only partially validated. Stand-alone biological modelling has attracted some attention from computer scientists. In particular certain biological phenomena such as biochemical networks appear to lend themselves to representation in formal schemes such as process calculi, opening the possibility for formal analysis and reasoning. A good sample of such work can be seen in [6]. Only a very small proportion of stand-alone models are accessible to other than their developers or are documented in a form other than the scientific papers in which they originally appeared.

Model Integration State-of-the-Art

It should be readily apparent from the discussion above that model integration will be very important for systems biology. Only recently however, has model integration received the attention that it deserves. In general the state-of-the-art is represented by ad-hoc, handcrafted, integration of stand-alone models and is characterised by tight coupling between these models.

The Systems Biology Workbench Project is a contribution to filling this gap. It comprises two distinct components: the Systems Biology Markup Language (SBML) [7] and the Systems Biology Workbench (SBW) [8]. SBML is an XML language for representing biochemical network models. It has largely been driven by a pragmatic concern to make the exchange of models across a range of popular modelling tools easier. SBML has achieved some success in this regard. SBW is a software framework to support interoperation among the heterogeneous tools and resources used in biological modelling by way of a relatively low-level message passing and brokering architecture. The SBW standard is not specifically tailored to biological modelling, but is instead a generic middleware solution. Though neither SBML nor SBW focus on model integration directly SBML provides a common framework for documenting a small range of models, which is an important first step towards model integration.

CellML [9] has been developed in parallel with the Physiome Project. It is an XML based language whose principal aim is the storage, exchange and ultimately reuse of biological models. CellML provides a high level ‘block diagram like’ representation scheme in which networks of models can be assembled and hierarchically composed. It makes significant use of the XML namespace mechanism to embed other languages such as Math ML. Some serious attention has been paid to descriptive metadata, though this remains a less developed aspect of the project. Unlike SBML, CellML explicitly attacks the model integration problem. Like SBML however, CellML can only encompass a limited range of models excluding, for example, discrete event systems. CellML is less widely used than the more pragmatically driven SBML.

Challenges

In order to map out the systems biology space more systematically, and to identify the computational challenges more precisely, we use a simple and very high-level information model [Figure 2]. It constitutes a metamodel for systems biology. We have used a very simple and stripped-down modelling convention: Entity-Relationship modelling.

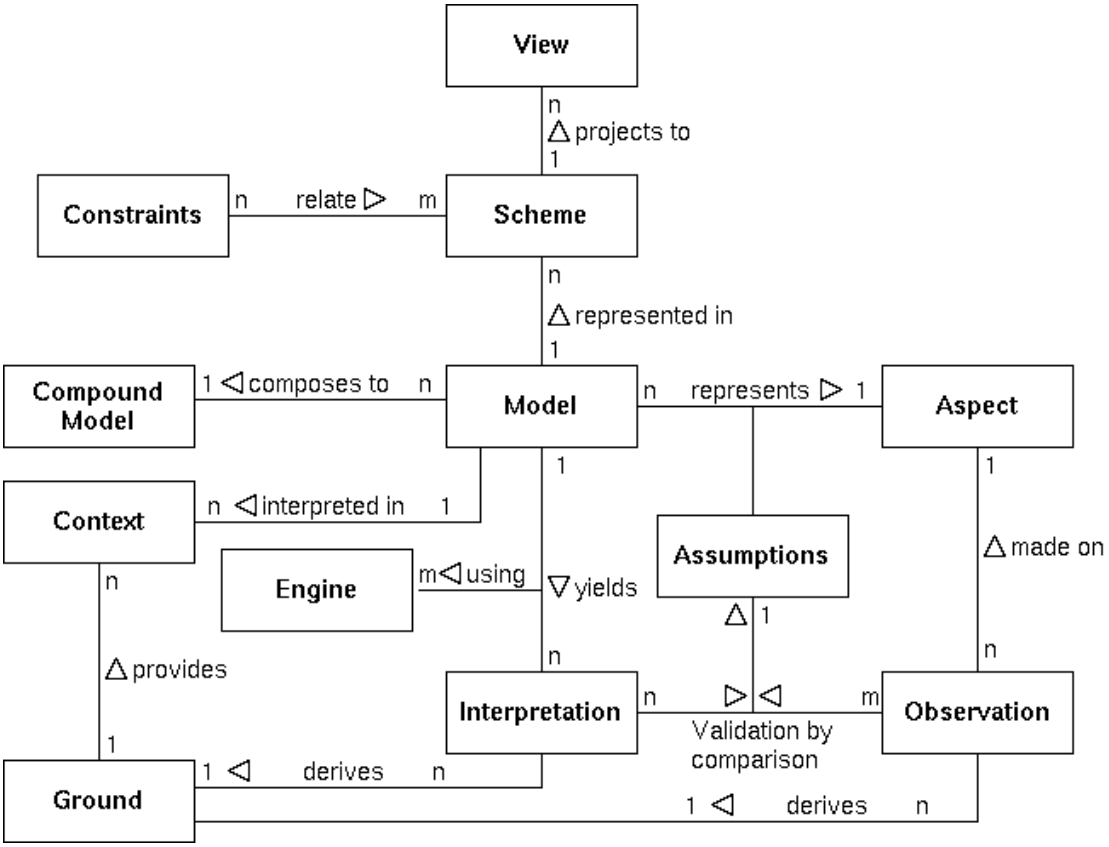

Figure 2: An Information Model for Systems Biology

The information model has three overlapping 'regions', each representing a key concern in systems biology: construction (basically the upper part of the model (model — compound model — scheme — constraints — view); analysis (model — context — engine — interpretation — ground); validation (model — aspect — observation — assumptions — interpretation).

We shall begin our discussion of this structure with model validation.

Models represent *Aspects*. We use the term aspect to denote a coherent set of properties or phenomena of biological interest. This is the anchor of the model in the real world.

The means by which a correspondence may be established is through an *ontology*, an explicit formal specification of how to represent the objects, concepts and other entities that are assumed to exist in the biological domain of discourse and the relationships that hold among them. The model and, as appropriate, elements of the model must then be linked to elements in the ontology.

The relationship between models and the aspects they purport to represent are conditioned or determined by *Assumptions*. Assumptions underpin model construction. They constitute the rationale for the model and must be precisely documented and connected to the model if it is to have meaning beyond the immediate use to which it has been put.

Experimental biologists make *Observations* on phenomena of biological interest. Classically, these are used to validate the *Interpretations* derived from models. Commonly however, the interpretations yielded by models themselves prompt further observations or, by comparison with observations, question the validity of the assumptions. The observations are documented in the scientific literature and in data resources associated with the experiments. The tie between descriptions of experiments, observations, experimental data, interpretations derived from models, and assumptions is one of the central challenges of systems biology. In short, systems biology cannot be seen independently of an information management framework that embraces a significant part of the experimental life sciences.

Scientific understanding is an inductive process, with a series of observations agreeing with a model each adding to our confidence that it is a good reflection of the system it purports to describe. Thus, validation is a more troubled concept than in some areas where modelling is used, since it is a matter of degree, rather than certainty. Refutation is in principle much simpler, but care must be taken here too, in deciding how best to modify the model to take into account a disagreement between a previous version and observation. Many believe that it is in these circumstances that modelling is at its most useful for developing scientific understanding. If we put our best scientific understanding into a model, and it does not fit the data, it suggests that our understanding is incomplete. This can be a powerful guide to new theories to try and new experiments to carry out.

Models, once instantiated, yield interpretations through a process of analysis. This may be dynamic, commonly simulation, or static, commonly a process of mathematical reasoning. The *Engine* that, abstractly, both encompasses and executes a model, determines the analytic process. The same model can, of course be analysed in many different ways and by different procedures. The engine thus conditions an interpretation. We are required to precisely specify the engine in order to anchor the interpretation. In short, it is not enough to define the model but also the means by which the model is used. Analysis may require significant computational resources.

Context is the data required to produce an instance of a model, though not all models require such data. It is the input to the model. A context must, of course, be soundly derived. This could be from observation, as in the straightforward case where experimental results provide a *Ground* for data supplied to a model. An alternative is the somewhat more complex case when the interpretations yielded by one model constitute the context for another model. From an informational standpoint, we need to track the contexts supplied to the model and associate them with the interpretations to which they correspond. We also need to track elements of context back through their grounds, because the integrity of the validation depends on them.

Model construction is, from a computer science standpoint, well-trodden ground. Models are constructed in different languages, or representation *Schemes*, each appropriate to the expression of, and reasoning about different sets of properties. No 'universal' language for systems biology can capture the many different phenomena we are interested in. The presentation of these schemes is through *Views* defined as projections on the underlying scheme. Modelling schemes are related to each other through *Constraints* that define what it means for models in these schemes to be consistent with each other. Most schemes aimed at modelling 'in-the-large' provide a compositional mechanism to permit models to be composed and larger-scale *Compound Models* to be constructed.

Just as this picture is familiar, the challenges it presents are well known. We must separately define and manage the views, the language definitions, and the constraints. We must provide means for checking the constraints. We must devise modelling schemes with sound compositional mechanisms. We must manage models that may not be consistent with each other, either across schemes or across scales. There is of course ample scope to extend the range of modelling schemes used in systems biology, and the extensive arsenal of formal modelling techniques that have been developed by computer scientists can be usefully employed here.

This overall picture is complex, but nevertheless excludes two key dimensions. Models may be produced in different versions over time and by different teams. There may be disagreements and contested observations. Models may be reused in different versions and configurations by different researchers. The result of the systems biology endeavour is thus unlikely to be a set of canonical models but rather a complex 'ecology' of models embedded within a framework in which debate and collaboration among contributors is enabled. Ultimately, our objective might include individualised models that take account of individual variations in physiology rather than generic models of biological phenomena.

Modelling The Liver

Having set out the challenges, the scale of the task facing the scientific community can be seen. It is necessary therefore to order and prioritise, in the form of a research agenda, the actions that need to be taken. As a first step we need additional convincing exemplars of systems biology of the general type of the heart model discussed above. Such examples will necessarily be restricted in scope and scale. Ideally however they will be more explicitly 'engineered' with some systematic modularity and separation of concerns among component models. These models can then act as test-beds for the broader conception of systems biology and for the information management frameworks that must accompany it.

The liver is the subject of a major UK research project funded by a Department of Trade & Industry 'Beacon' scheme that supports 'high adventure science' with the possibility of advances that have significant industrial potential. The aim of the project is to produce a physiological model of the liver that is integrated across scales [10]. The liver has been selected as a good exemplar of systems biology: it is medically important and structurally relatively homogeneous. It is also challenging – the liver is primarily a chemical system, where the heart is electromechanical. Electromechanical systems have a long history of quantitative description, and modelling in that area is thus comparatively advanced. There are also a number of ongoing efforts to build 'in vitro' livers, that is artificial livers that can be used while patients who have suffered liver damage are recovering. Effective models of the liver could be used to understand and overcome some of the problems experienced by builders of such livers. The project brings together physiologists and experimental life scientists, engineers with expertise in systems modelling, applied mathematicians with an interest in integrating models across differing temporal and spatial scales, and, of course, computer scientists who can put the information management and computational infrastructure in place.

The liver [Figure 3] has three principal functions: it stores materials to be released into the blood stream when needed; it synthesizes proteins and peptides from amino acids; and, it detoxifies the system by breaking down harmful materials such as alcohol, which are then excreted.

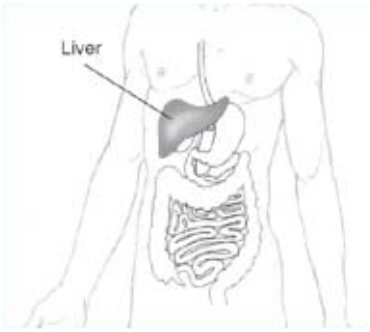

Figure 3: The Liver (US National Diabetes Information Clearinghouse)

To illustrate current work on systems biology let us look in some more detail at one of the better understood examples of the first of these functions: glucose release from the hepatocyte (liver cell) in response to circulating adrenaline or glucagon. Adrenaline is part of the classical 'fight or flight' response to stress. Glucagon is part of the homeostatic control of blood glucose. Both these systems are compromised in diabetes, when cellular uptake of glucose driven by insulin is defective. Both adrenaline and glucagon activate the same intracellular mechanisms: the hormone, circulating in the blood stream, binds to specific receptors on the membrane of the hepatocyte. As a result, ion channels (specialised protein molecules that allow specific ions to enter or leave cells) open in the membrane. Calcium enters the cell through the ion channels and the cytoplasmic (cellular material within the cell membrane outside the nucleus) concentration of calcium rises. Binding of hormone to receptor simultaneously activates linked G-proteins and initiates a chain reaction within the cell that also causes cytoplasmic calcium to rise by releasing calcium from stores within the cell. At different concentrations, calcium both stimulates and inhibits calcium release from stores, and consequently cytoplasmic calcium oscillates. The increase in calcium also mobilizes glucose release from glycogen (the stored form of glucose), which leaves the cell on glucose transporters.

This pared down description amply illustrates the complexity of the dynamic relationships involved in relatively straightforward physiological processes.

Models of each of these sub-processes, such as G-protein activation or cytoplasmic calcium oscillation, may be constructed in isolation. Typically these may be modelled as ordinary differential equations (ODEs) though certain processes appear to lend themselves to discrete event modelling. The processes have, in this case, been well studied experimentally and the parameters, that constitute the context, can be related systematically to values in the literature. Ideally this should be done by way of a mediating ontology. There are a number of significant projects constructing such ontologies for human physiology. An example is the Digital Anatomist Foundational Model [11]. The richer ontologies developed for genetic and bioinformatic work, such as the Gene Ontology [12], can also be useful for cell physiological work. It is, of course, necessary to look carefully at the reliability of the experimental data when selecting the parameters to use with the model.

Assuming homogeneous models of the sub-processes we can connect these together to build a detailed model of the entire network. Representational heterogeneity naturally makes this more difficult. The resulting model can be investigated numerically by varying its context.

Alongside this model we can build a 'simplified' or gross model. Rather than the more complex behaviours built into the models of ion channel opening, protein activation etc.

we assume these behave as perfect switches to make the system piecewise linear. The simplified system is biologically unrealistic, and many features, such as the shape or period of oscillations, are not preserved. Some, however, are, and the simplified model allows us to use algebraic analysis, facilitating the development of understanding about the system. For example, we can learn how specific features of the system's behaviour are controlled by certain elements of the context. Even in the absence of analytical results, a model that is simple enough to hold in the human mind in one go is a useful tool for understanding, and as a comparator to the fuller, more unwieldy, model.

Both the detailed and simplified models are constructed and analysed within standard tools for scientific modelling which need to be wrapped to support model integration. They must be connected to standard scientific visualisations, such as graphs, and we could readily envisage more sophisticated animated views.

We could (and intend to) take modelling of this system a long way further. An immediate extension is to incorporate the homeostatic activation of glucose release through glucagon receptors. We could, for example, build models of gap junctions, which are constructed from membrane-inserted proteins ('connexins') that bridge the space between cells and provide direct channels that allow the cytoplasm of one cell to communicate with that of adjacent cells. By this means we could link more than one cell, scaling-up to multicellular models. Another approach to the scaling issue would be to consider the effects that signalling molecules have on gene expression by acting as transcription factors (proteins that bind to regulatory regions), thus moving down to the molecular machinery.

Modelling Strategy

As we have suggested, it is not feasible to try to represent all aspects of a biological system in the smallest conceivable detail, even when data is available. We cannot, and do not need to, recreate the world as an isomorphic in-silico image of itself. The art of systems biology will therefore largely be driven by 'judicious simplification'. This is particularly true when trying to link different processes at different scales (spatial and/or temporal); for example, gene and protein networks. Some simplifications may be appropriate for some questions one might ask, and others for others. For example, in order to represent biochemical networks that involve many different proteins, we may represent the interactions between proteins as simple stimulus-response functions. Alternatively, we may choose to focus on a few proteins, and model the extremely complex transformational processes between them in great detail.

Simplification has at least three facets:

A) Choice of a modelling scheme. This must provide sufficient descriptive fidelity, flexibility when linking to other models, contextualisation in terms of known

(or obtainable) data, and reasonable ease of interpretation.

B) Choice of level of detail within the given representation. How many 'links' in a signalling pathway really need to be explicitly represented? Does space have to be explicitly modelled, and if so how (there are many ways)? What is the dominant time scale?

C) Determining sensitivity. A simplification scheme is not much use unless its context and interpretation are 'robust' (in some poorly defined, because poorly understood, sense). If a model turns out to be a delicate flower, then, more than likely, important elements of 'backbone', which give robustness to the real biological system, have been omitted. It is, of course, always possible that the real system also is sensitive, and this should not be ignored.

Some of these issues may be clarified by thinking about the interpretation obtained from the simplified model of calcium oscillations (with square waves) discussed above. It is important to realise that one value of this model lies in the fact that it is at the extreme end of a continuum of models of the same general type (with Hill function response functions), all of which behave in a qualitatively similar manner. That means we have a continuum of 'possible worlds', only at most one of which is a 'true' representation of the real world. The fact that all these models behave qualitatively in the same way tells us that, in some sense, the detail of the real-world response may be incidental: there may be recognisable (potentially real) worlds in which calcium oscillations are different. Thus, some perfectly feasible (intergalactic!) creature may have square calcium waves. The fact that we humans have the shape of wave we do is therefore attributable to some kind of 'fine tuning'. We must determine how strenuously we should chase this kind of fine-scale effect, rather than be content with more robust, qualitative phenomena. It can be a difficult, fraught question, deciding which behaviours a model must reproduce. This is, in our experience, an area in which the instincts of researchers from different backgrounds can disagree strongly.

In this context 'function' can be an important guide for model construction and interpretation. That is, we know (roughly) what a liver is 'for'. However, this may not always be the case. At the fine grain in biological systems we can observe phenomena whose function we do not understand. In a deep sense they may not be *for* anything — there is no logic to evolution. This can make modelling much more problematic, because we don't really know what we are aiming for a model to do. If we do not know which phenomena are central and which are incidental, it is extremely difficult to assess the validity of any model.

It is important to realise that in many cases, the 'experiment' not only has to be performed in the laboratory, but also itself modelled. That is, we must model not just the physiological process, but also the experimental protocol. For example, we have been much exercised with an experiment that, when taken at face value, seems to refute an assumption underpinning a certain model of protein production by cells.

This, however, turns out only to be true if the model is interpreted in the most naive way. There are more sophisticated interpretations (involving the explicit representation of stochastic effects) in terms of which the laboratory experiment may be analysed and its result predicted, but which *are* compatible with the original hypothesis. It is not always clear just what an experiment does, and does not tell you about a model.

Model Integration Strategy

The relationships between models in different schemes are thought of, in our framework, in terms of constraints that define what it means for those models, when placed in conjunction, to be consistent with each other. Expressing these constraints, or understanding how the models relate, poses many difficulties when the models to be integrated are of very different kinds. Figure 4 below, shows a simple taxonomic framework that contrasts modelling schemes based on different principles. Particular problems arise when models are stochastic, or some models are formulated as discrete-time systems and some as continuous-time systems. Our strategy assumes that it is preferable to design coherent collections of models rather than struggle to integrate fundamentally incompatible schemes. What these should be, and how they are structured remain open questions.

Deterministic	----------	Stochastic
Compartmental variables	----------	Individual or functional
Spatially homogeneous	----------	Spatially explicit
Uniform time scale	----------	Separated time scales
Single scale entities	----------	Cross-scale entities

Figure 4: A Taxonomic Framework for Modelling Schemes

Where Now?

Unlike projects to map genomes there is no clear end-point for systems biology, although we can, perhaps, identify some important staging posts. Models that provide 'thin' vertical slices across scales are one such. Our models of glucose release in the hepatocyte are already very close to cross scale integration from gene expression through to multicellular responses. As another example, genes for aquaporins (membrane water channels), which control the movement of water into and out of cells have been identified in the past 5 years. Fluid transport is a key part of liver physiology. Moving from the gene through models of aquaporins to bile flow would be a significant achievement. Another staging post is the development of models that are approved for drug testing, perhaps in place of animal models, and that satisfy the strict requirements of validity, reliability, transparency and traceability. The establishment of

global ‘collaboratories’ in which models can be exchanged, reviewed and analysed will also be important. Finally, when we can dependably diagnose health issues and identify novel treatments using our models, systems biology will have come of age.

References

[1] D. Noble, "Modelling the heart: from genes to cells to the whole organ," *Science*, vol. 295, pp. 1678-1682, 2002.

[2] D. Noble and T. Colatsky, "A Return To Rational Drug Discovery: computer-based models of cells, organs and systems in drug target identification," *Emerging Therapeutic Targets*, vol. 4, pp. 39-49, 2000.

[3] A. Hodgkin and A. Huxley, "A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve," *Journal of Physiology*, vol. 117, pp. 500-544, 1952.

[4] Description of the Physiome Project, *Physiome Project*, [online] 2003, http://www.physiome.org (Accessed: 2004 January)

[5] M. Levin, C. Morton-Firth, W. Abouhamad, R. Bourret, and D. Bray, "Origins of Individual Swimming Behavior in Bacteria," *Biophysical Journal*, vol. 74, pp. 175-181, 1998.

[6] C. Priami [Ed.], "Computational Methods in Systems Biology," Proceedings of the First International Workshop on Computational Methods in Systems Biology 2003, Roverto, Italy, 2003.

[7] SBML: Systems Biology Markup Language, *sbml.org - The Home Site for the Systems Biology Markup Language*, [online] 2003, http://www.sbml.org (Accessed: 2004 January)

[8] The Systems Biology Workbench, *The Systems Biology Workbench*, [online] 2004, http://sbw.sourceforge.net (Accessed: 2004 January)

[9] What is CellML, *The CellML.org Website*, [online] 2001, http://www.cellml.org/public/about/what_is_cellml.html (Accessed: January 2004)

[10] The UCL Beacon Project, *The UCL Beacon Project*, [online] 2003, http://grid.ucl.ac.uk/biobeacon/php/index.php (Accessed: January 2004)

[11] C. Rosse and J. Mejino, "A Reference Ontology for Bioinformatics: The Foundational Model of Anatomy," *Journal of Biomedical Informatics*, 2003.

[12] Gene Ontology Consortium, *Gene Ontology*, [online] 2004, http://www.geneontology.org (Accessed: January 2004)