

Systems Biology Meets Neuroinformatics

Neuroinformatics has been defined by the Organisation for Economic Cooperation and Development (OECD) as "...combining neuroscience and informatics research to develop and apply the advanced tools and approaches that are essential for major advances in understanding the structure and function of the brain". Systems biologists will detect a very strong parallel here with their own discipline. It is not surprising, therefore, that the neuroinformatics and systems biology communities are beginning to form close links.

The UK Neuroinformatics Network has been established with funding from the EPSRC, MRC and BBSRC to foster collaborations between clinical and basic scientists involved in the use of informatics to study brain and behaviour. It brings together the UK neuroinformatics community, and represents it in discussions with scientists in related disciplines, commercial partners, funding agencies and the public at large. The focus of the Network on establishing links with related disciplines was emphasised strongly in a recent workshop it organised. Held in Edinburgh at the end of April 2006, the Systems Biology Meets Neuroinformatics Workshop was co-sponsored by the Edinburgh Centre for Bioinformatics (<http://www.bioinformatics.ed.ac.uk>) and the Scottish Bioinformatics Forum.

The workshop format was chosen to encourage close and detailed discussions between researchers in the two disciplines. The number of participants was limited to around 50; the large majority of delegates were established academic researchers although industry and the funding

agencies were also represented. Much of the meeting was spent in breakout groups, with delegates discussing issues relating to the development of the two disciplines and drawing out suggestions for topics that should be addressed jointly. There were also five presentations: two from leading researchers in each area, and a summary lecture from **Seth Grant** of the Sanger Institute in Cambridge.

Systems Biology

The first plenary lecture in the systems biology strand was given by **Hans Westerhoff**, who now divides his time between the Free University of Amsterdam and the Centre for Integrative Systems Biology at the University of Manchester, UK. His talk on the modelling of metabolic and signal transduction pathways and networks – which is a crucially important part of any attempt to model brain function at a molecular level – focussed on the Silicon Cell programme (see <http://www.siliconcell.net>). This defines a silicon cell as "a precise replica of (part of) a living cell", as a computer model using equations and parameters based, as far as possible, on published experimental data.

Westerhoff described this approach to modelling metabolism and signal transduction using yeast as a model system. Even yeast cells can, in a very simplistic manner, be "trained"; if they are given pulses of glucose at regular intervals they begin to show synchronous oscillations in their metabolism, and these oscillations continue even when the glucose pulses stop. This "learning" must depend on gene expression, since adding an inhibitor of protein synthesis will prevent it. Westerhoff set out to discover the principles underlying this oscillation and other time-dependent patterns of gene expression using the principle of metabolic control analysis, a method pioneered at Edinburgh by Henrik Kacser. Modelling the signal transduction cascade induced by

EGF led to the unexpected deduction that the amplitude of the response depended equally on the concentration of kinases and phosphatases, whereas its duration depended much more on phosphatases. "No-one had any idea that phosphatases were more important than kinases for determining the duration of a signal response", he said. "This is an example of the utility of systems biology in determining general biological principles."

He applied a similar methodology to *Escherichia coli*, and showed that bacterium could "learn" to select a metabolic pathway for assimilating ammonia depending on the amount available. This "molecular neural network" might be seen as a primitive type of memory, even in a unicellular organism that self-evidently lacks a brain. Understanding processes like these may give us insights into the functioning of memory in much more complex organisms, including ourselves.

Nicolas Le Novère, from the European Bioinformatics Institute, Cambridge, UK, gave the second presentation. His talk focused on the need for common standards and models in what he termed "systems neurobiology". The principles he presented, however, would seem equally applicable to any other area of systems biology. Mathematical models of biological systems should be reusable for three reasons. Biologists who use models need to be able to change them without necessarily learning to program; "experts" need to avoid "reinventing the wheel" when modifying models; and the various components of a model system, which may be set up for simulations at very different spatial and temporal scales, need to communicate with each other. In addition, models should be what he termed both "computer-edible" and "human-edible", and they must be published and made readily available.

Systems biology markup language (SBML; <http://www.sbml.org>) is an

internationally supported and widely used “computer-edible” standard for encoding systems biology models, based on the generic XML standard. SBML is scale free, in that it can be used equally for models at the cell, organ and organism level; it can describe reactions and cellular processes as well as objects; and it is linked to and can be used with XML schema in related biological areas, including, in neurobiology, NeuroML and BrainML. The use of SBML or a similar standard is a good first step towards producing a model that is fully reproducible, but it is not sufficient.

In order to make a systems biology model “human-edible” it is necessary for it to be presented and described using clear and unambiguous terminology. Many branches of life science have introduced ontologies to standardise terminology; systems biology is no exception to this. The Systems Biology Ontology (SBO) has so far evolved four linked vocabularies: a taxonomy of reaction participants (e.g. “catalyst”); vocabularies for parameters and laws in quantitative models (e.g. “Michaelis constant” and “mass action kinetics” respectively); and a list of modeling frameworks (e.g. “continuous modeling”). The Systems Biology Graphical Notation (SBGN) project (<http://www.sbgn.org>), working along similar lines, aims to help standardise a graphical notation for computational models in systems biology, adding rigor and consistency to the often *ad hoc* diagrams in research papers in this area.

Neuroinformatics

Upinder Bhalla from the Tata Institute of Fundamental Research, Bangalore, India, began the first neuroinformatics presentation by asking the question “Where is memory?” The most logical answer to this complex question is to locate memory in the brain, the neurological circuit, the neuron and the synapse. Individual memory-related events

may be modelled at all these spatial levels, and it is only by integrating these levels that we will be able to understand how they occur.

It is fairly well known that learning and memory is more effective when stimuli are spaced in time, and some researchers have identified an optimum spacing of 5–15 minutes in experiments with rats and mice. This is consistent with the involvement of signal transduction pathways in this type of pattern recognition. Bhalla and his colleague Sriram Ajay observed that levels of the kinase ERK-II peaked in hippocampal slices about 10 minutes after stimulation¹. Simulations of calcium flux and signal transduction in the same system, using a biophysical model based on one first introduced by Traub in 1991, were able to reproduce the dominance of this apparently physiologically relevant timescale.

Kevan Martin from the Institute of Neuroinformatics in Zurich, Switzerland – one of the first institutions to use that name – is applying Francis Crick’s thesis that “[to] understand the function of a complex system, study its structure and knowledge of its function will follow automatically” to neurobiology, using studies of neocortex structure to shed light on neurological processing in the brain. A large majority of fibres in the white matter of the brain connect one area of cortex to another; he is looking for the rules that constrain which cortical areas connect together. Models of the structure of the mammalian primary visual cortex go back to the work of Wiesel and colleagues in the early 1980s². This cortex has been divided into levels; Martin has found that the neurons in each level make synapses with those in the other levels in a characteristic pattern. This pattern is being encoded into a series of rules constraining synaptic patterns that can be tested in wider and more complex models. It is now technically feasible to model synapses within a 1mm x 1mm section of mouse brain, although any reconstruction of the complete mouse

brain is still likely to be many years away.

Group Discussions

Participants were divided into groups that were evenly balanced between the “systems biology” and “neuroinformatics” areas and between established and junior researchers, and invited to discuss a number of issues relating to the two fields and the interaction between them.

In the first session, groups were asked to define the two areas, give examples of specific research questions they address and identify the differences and potential overlaps between them. Participants were in agreement that systems biology “examines the interaction of biological components. . . to understand and predict emergent properties of systems”, and that neuroinformatics concerns modelling of processes in the nervous system and the storage, analysis and visualisation of neuroscience data. Thus, from these definitions alone, neuroinformatics might be seen as a subset of systems biology. However, participants also thought that the very complexity of neuronal networks and systems made it more useful to consider the disciplines as overlapping. One group showed this visually, using a Venn diagram with two circles overlapping by over 75%. They suggested that the only area of neuroinformatics that did not overlap with systems biology was that “small” research field concerning information processing and cognition in the brain. Common research interests were identified as – self-evidently – molecular signaling in neurons, but also the more general analysis of neural networks and their properties. Several groups identified common needs between the two communities, including the need for richer data from which to build models on many scales and for better visualisation and informatics tools.

In the second session, groups were asked to identify research questions

which would benefit particularly from a common approach by the two communities, and to identify opportunities for developing methodologies that would benefit both fields and common training needs. The research questions most often identified were basic methodological ones rather than specific biological questions, with groups focusing on the need to develop standards for data sharing and ontologies. With researchers in both disciplines committed to the development of open standards, there are growing opportunities to collaborate in developing software and modelling languages. It would be useful, for example, for SBML to include some of the definitions and parameters required for the development of neuroinformatics models.

One group raised the idea of generating a “connectivity map” of the brain, including molecular level data in a model of the neuronal network. Visualisation was mentioned again, as both a methodology and a research question. And, when it came to training, the groups were all but unanimous: participants stressed the need for “more mathematics” for life science students at all levels of education, from high school through to postdoctoral. This can be addressed, at least at the postgraduate level and above, by further integrating training between university departments of biological and computational sciences. In the Edinburgh neuroscience course, physical science

PhD students are embedded into an intensive postgraduate neuroscience course in their first year. The establishment of, now, seven systems biology centres at UK universities including one at Edinburgh³ must encourage similar initiatives in this field.

Systems Biology and Neuroinformatics: The Future

Seth Grant of the Sanger Institute near Cambridge was given the ambitious task of addressing the future of the two disciplines. He began with an anecdote. “If you ask ‘average people in the street’ what makes us human, they will not be particularly interested in our livers, kidneys or other organs: they will want to talk about the brain”. And, he said, the question of whether it was possible to predict brain function and behaviour from a genome was now at least tractable. However, the problem had to be addressed at different scales. He focuses on the molecular end of the continuum, using systems biology to approach the most fundamental unit of computation in the nervous system: the synapse.

We now know what a synapse is “made of” at the molecular level – that is, we have the “parts list” and can begin to ask questions about how those parts are integrated. Grant’s group is part of the Genes to Cognition consortium, supported by the Wellcome Trust, which is studying the effect of particular human and mouse genes on the synapse mechanism using a variety of

techniques including mouse knockout models, clinical genetics and electrophysiology. Previous studies, between 1992 and 2005, have identified a total of 134 different mouse phenotypes affecting synaptic plasticity from 190 single gene knockouts. However, with complex proteomics techniques, it is now possible to identify over 1000 separate proteins that make up the “post synaptic proteome”. Grant, working with Douglas Armstrong and Andrew Pocklington at Edinburgh University, has published the first systems biology studies of the synapse⁴. The group mapped protein interaction networks on to genetic phenotypes and proposed a general architecture for the organisation of the synapse proteome, enabling prediction of physiological properties and disease phenotypes.

In Conclusion

It is already clear that some of the exquisite structure described by Grant and others is shared by molecules and complexes in other signaling pathways, and even single cells have some limited ability to “learn”. Elucidating the level or levels of complexity at which brain functions emerge is one of the “grand challenges” that will need a fuller integration of the techniques thought of as “neuroinformatics” with those of “systems biology”. Hopefully, collaborations begun and fostered at the Edinburgh workshop will aid this integration.