

Physics of Life Network Workshop on Biocomputation

**Organisers: Sarah Harris (Leeds), Susan Stepney and Agnes Noy (York),
Viv Kendon and Dominic Horsman (Durham)**

The ability of living systems to store and process complex information in a manner that is responsive to the external environment has inspired the development of the multidisciplinary field of biocomputation, which aims to build computational devices using biomolecular systems. This workshop brought together researchers with interest in biocomputation with specialists in gene regulation, DNA mechanics, bioinformatics and computer science. The themes that emerged most strongly from the talks and discussion sessions at the meeting were the potential applications of biocomputation, the requirements for new computational languages and models, and the most challenging open questions in the field.

Biocomputation: Its future potential. Biology provides a unique example of a naturally occurring information processing system. If we can understand how this is achieved biological information processing, then we should be able to mimic it synthetically. It is this challenge that defines the field of biocomputation. Three distinct strands in the field were articulated by Dr. Katherine Dunn during the workshop: i) Computing FOR biology – e.g. computer models of biological systems ii) Computing LIKE biology – e.g. computational methods or algorithms that mimic biological systems, such as neural networks or evolutionary algorithms and iii) Computing BY biology – e.g. using biological systems to build computers. This workshop focused on the third of these themes.

To be most useful, biocomputation should exploit the many unique properties of biological systems compared to established computer hardware. Firstly, biocomputational devices have the potential to interact directly with existing biological systems in a manner that is not possible with non-organic technologies. In addition, the stochastic nature of biological interactions allows logical operations within a population of cells or synthetic biomolecular computational compartments to be performed in a probabilistic, rather than deterministic manner, which may offer opportunities for unique types of calculation that are impossible with existing solid state computational devices. Furthermore, biological systems can be inherently tolerant to noise, they can produce emergent behaviour from a set of simple rules and they can provide analogue computation, which is complementary to the digital logical operations performed by conventional computing. Biological systems are also inherently “low power”, which could solve one of the major global problems faced by the expansion of computing infrastructure on a global scale computing industry, while biological systems also have the ability of biological systems to self-replicate, evolve and adapt, and they can adapt to changes in their environment, which provides both opportunities and challenges for the design of computing living systems that compute.

Languages and models for biocomputation: How do we articulate the complex network of interactions between computing biological components? The explosion of biological online databases increasingly requires that this knowledge be represented in a standardised form, so that research communities, and now often increasingly data searching algorithms, can locate and process this information efficiently. For example, in the field of synthetic biology, the Internationally Genetically Engineered Machine (iGEM) project provide an online Registry of Standard Biological Parts, which is a standardised repository of genetic components that can be combined to build synthetic biology devices using a predefined set of assembly rules. One of the challenges facing the biocomputation community is how to establish the appropriate set of rules for combining these and other components in practice, as current attempts frequently do not produce the expected behaviour once assembled into a larger system. For DNA-based computational engines, this may be because the context and location of a given DNA sequence affects its output how it is processed. Understanding how the physical status of genes affects their regulation will be key to determining a successful set of design rules for biocomputational devices based on gene expression. Equivalent dependence on the physical location and environment of proteins in regulatory networks is also a limiting factor in understanding the function of cell signalling and metabolic pathways.

Obtaining sufficient physical insight that the output from networked biological components can be predicted opens up the possibility of a semantic language for biology, in which the computational output from a given

design can be demonstrated to have particular properties. Such a tool would be as valuable to our understanding of existing organisms as it would to the engineering of synthetic counterparts, because it would offer a method to integrate the individual biological components to form a holistic model of the overall outcome of the network, in an analogous manner to the compilation of computer code to generate an executable program. The invention of computational languages for synthetic and systems biology is an active area of community driven research. The Synthetic Biology Open Language (SBOL) for example, provides a vocabulary of schematic glyphs to represent genetic designs, whereas the Systems Biology Markup Language (SBML) is a machine-readable semantic language for model building and analysis of signalling and metabolic pathways and gene regulatory networks.

Understanding and applications: What are the barriers to providing biocomputing technologies?

One of the ultimate tests of understanding is our ability to engineer. An important role of synthetic biology is that it strongly tests our knowledge of living systems, and so exposes the areas where further research is most needed. As yet, the use of biology-based computational devices is limited to research laboratories, and no technologies are in widespread use. The barriers are primarily in understanding the complexity of biological systems with sufficient fidelity that we can create synthetic versions of them that are both sufficiently complex that they can be used for computation, but which still behave in a predictable manner. This is an exciting research challenge that should be a focus for multidisciplinary research at the Life Science Interface.

Presentations

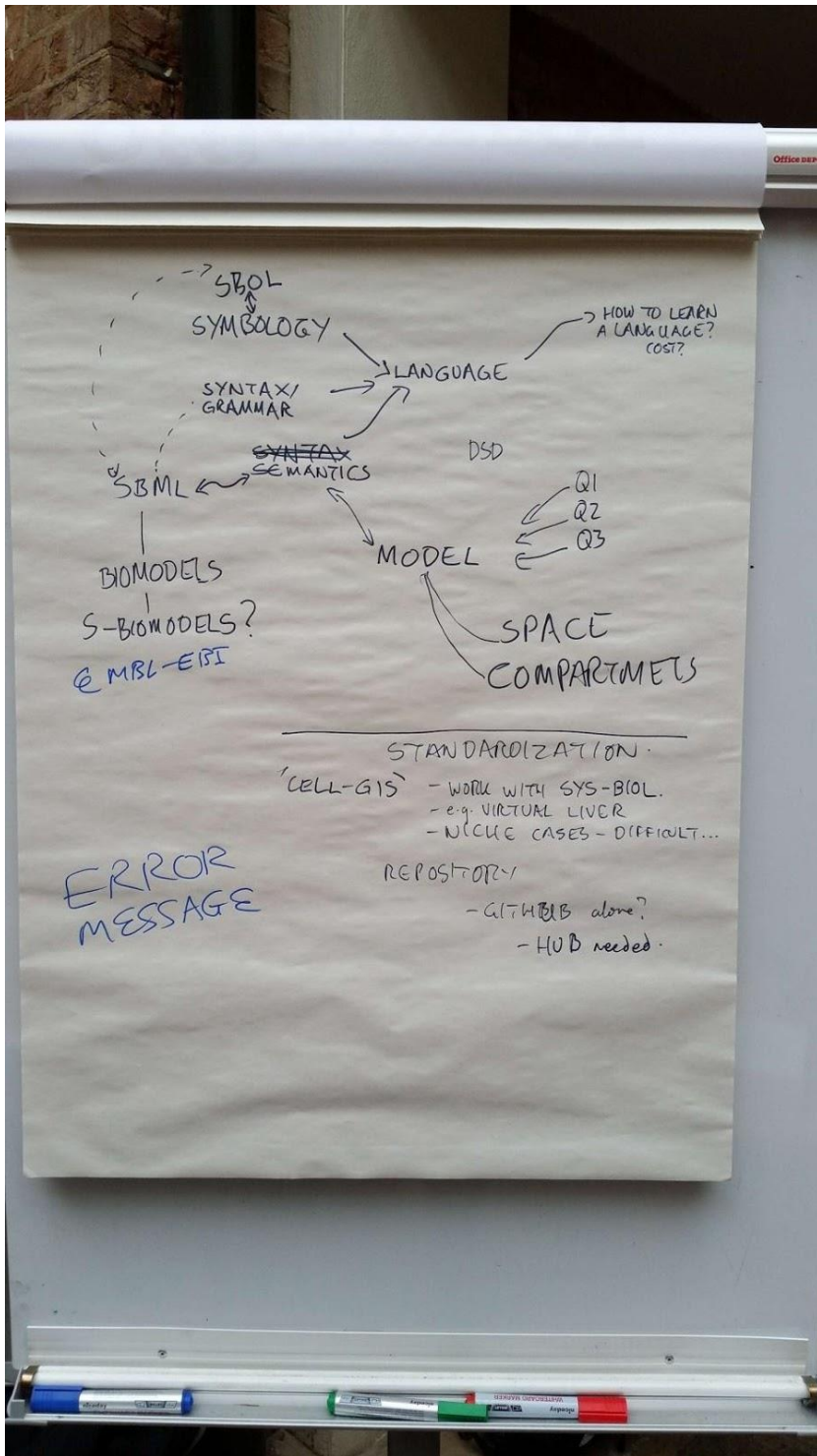
- Overview of the Current State of the Art in Biocomputation: Successes and Challenges. Andrew Turberfield.
- In Vivo Synthetic Computation. Rob Bradley.
- How do living organisms process information? Charlie Dorman and Sean Colloms.
- Measurement and Detection: What are the current technical limits to determining the output from biocomputation? Mark Leake and Massa Shoura
- Languages and models for biocomputation: Katherine Dunn and Simon Hickenbotham
- Applications of biocomputation: Angel Goni-Moreno

Photos

Conference photo:



Photo of one of the flip chart pages:



Relevant web links:

iGEM: http://igem.org/Main_Page

SBOL: <http://sbolstandard.org/>

SBML: http://sbml.org/Main_Page