From "Cells as Computation" to "Cells as Apps"

Andrea Bracciali*, Enrico Cataldo**, Claudio Felicioli***, Roberto Marangoni**, Pasquale Stano****
* Stirling University, UK / ** University of Pisa, Italy
*** Fabnami, Switzerland / **** University of Roma Tre, Italy

Biology is being understood also as a quantitative and computational science. Distributed information processing and the quantitative dynamics of massively distributed molecular systems are key when investigating life systems. Current breakthroughs in synthetic biology bring into play the possibility of programming such computational biological machineries. We reflect on these recent ideas and the perspectives they open. As usually happens in interesting and developing fields, open challenges largely outnumber answers.

The results of modern genomics and other –omics disciplines, greatly improved our knowledge about the biochemical and functional network that regulates the cellular life, and we have strong insights on some of the most important components of the network. Many control structures have been identified (e.g., power–law in genetic networks, [6]), and important links between the network topology and functional aspects clarified (Flux Balance Analysis [7]). It has become evident that some solutions found by the evolution have optimized not only the physical–chemical properties of several molecules but also their computational properties. Recent contributions have clarified the information processing nature of protein interaction within a living cell [1], the need of a systemic approach for understanding cell functioning [2], and the suitability of computation as understood in computer science, as an interpretation key [3], leading to the paradigm *"cell as computation"*. Novel research directions have spawned from the approach, e.g. DNA–based technology for implementing the computational core of sensing/control processes [4], RNA–based logical gates as the basis for the development of "in vivo logic processing" [13].

Such recent progresses bring into play the *programming* aspects of computation. The theoretical possibility to control the cellular molecular machinery leads us to imagine programmable devices based on biological cells, a new frontier in biology, where computer science, engineer, and physics work together to design new organism or protocells able to perform targeted biological actions.

The limits and long–term implications of such a scenario are fully to be understood. One of the reasons why synthetic biology [8] has become very popular is that it promises endless possibilities of manipulating organisms for achieving a predefined function, although to date only microorganisms have been approached. Current and envisaged applications regard the production of biofuel, chemical intermediates for pharma–chem industry, destroying pollutants, drug development and delivery, and diagnosis and therapeutic tasks.

Together with modifications of extant organisms [9], one promising approach for the bottom–up design and engineering of synthetic systems from their constitutive components is the design of standardised biological parts in a systematic and rational manner (the *semi–synthetic approach* [10]) where minimal

sets of biological macromolecules are encapsulated inside liposomes. Which computational power do semi-synthetic cells have? They share with natural cells the capability of "chemical computing", i.e., processing information by manipulating chemical "signals". Chemical computing comprises molecular recognition, transformation, control activities, resounding constructs in traditional programming languages. Chemical computation can be exploited to build micro-machines that are capable of unconventional computation [11].

How biologically embedded computation relates to classical Turing-universal computation has clearly attracted interest, e.g. [5] for bio-chemical computing, and [14, 15] on the universality of DNA computing by reduction to grammars systems and specific classes of cellular automata. Several results prove universality of biocomputation, others strive to define suitable theoretical model for what appears to be more suitably described as a reactive and distributed system, e.g. parallel cellular automata, than a conventional Turing machine. Moreover, probabilistic and stochastic phenomena play an essential part in bio computation, calling for models that account for that, such as Markov chain based models, which naturally link to the Gillespie approach for biochemistry [16]. It is worth observing how bio-embodied computation heavily depends on the "hardware" (although this seems not to be a suitable term) on which it is executed, differently from Turing computation that is not bound to any specific executor, such as the von Neumann architecture.

The fact that, quite differently from classical programming, the execution environment is largely uncontrollable and currently poorly understood represents one of the main limitations and difficulties of a full-fledged synthetic programming. Modularity, one of the desirable property when designing complex systems, often does not hold, since the behaviour of the whole system is often a possibly stochastic *emergent* property, which can be observed only at system level and can not be easily understood from the collection of the components considered in isolation. Biological computation observed in living organisms is not always clearly recognizable as classical computation. Which are the missing constructs of a programming language for biological computation? Are differences due to the different ways in which "the program is being developed", i.e. evolution vs. human design?

Looking further ahead, which are the implications when scaling to organs and organisms? For instance, will we be able to program the healing or well-being of an organ? And of an organism? Interestingly, can we program the brain as we aim to program other organs? Although the latter may sound a bit visionary, there has been a lot of progress from the publication of the Hodgkin-Huxley model of the neuron, and models of action potentials, inter-neuronal communication, dendritic spines and synaptic boutons dynamics, neuromodulation, plasticity and development have been proposed [18]. Neuronal activity can be measured in the living brain and correlated with ongoing behavior. To overcome correlation and investigate causal impact of neuronal activity to behavior, new techniques have been developed to take control of single neurons in a living brain with high spatiotemporal resolution: optogenetics let us control neuronal activity in-vivo by using photons, and optogenetic actuators can be placed in specific cell types gaining high spatiotemporal resolution. To program animal behaviour by manipulating neuronal activity has been proved possible, but major improvements in both models and techniques are needed [17, 19].

Synaptic plasticity is understood to be the mechanism underlying learning and memory,

and several insights have been obtained elucidating where memory is stored, how many types of memory exist and what are the molecular basis and the structural mechanisms of some forms of memory. Biological memory storage is the result of the interaction of several processes. During the retrieval, update and integration of a given memory with other memories, it is possible to obtain destabilization and restabilization of memory. In some forms of memory, the retrieval of long–term traces can yields a transient destabilization of the traces and a persistent change of them. A complex molecular machinery is involved in such memory processes [12]. Roughly speaking, one could see these processes as a molecular foundation of psychology, which manipulates and heals disturbing memories through recalling them. Can these molecular machines be programmed? Which are the implications?

References

[1] D. Bray, "Protein molecules as computational elements in living cells", *Nature* 376, 1995, pp. 307-312.

[2] H. Kitano, "Systems Biology: A Brief Overview", *Science* 1, 295 (5560), 2002, pp. 1662-4.

[3] A. Regev, E. Shapiro "Cells as computation", *Nature* 26, 419(6905), 2002, pp. 343.

[4] Y.-J. Chen, N. Dalchau, N. Srinivas et al., "Programmable chemical controllers made from DNA", *Nature Nanotechnology* 8, 2005, pp. 755-762.

[5] L. Cardelli, G. Zavattaro, "Turing universality of the Biochemical Ground Form", *Mathematical Structures in Computer Science* 20, 2010, pp. 45-73.

[6] F.M. Alakwaa, "Modeling of Gene Regulatory Networks: A Literature Review", *Journal of Computing System Biology* 1(1), 2014, pp. 1.

[7] C.P. Long, M.R. Antoniewicz, "Metabolic flux analysis of Escherichia coli knockouts: lessons from the Keio collection and future outlook", *Current Opinion in Biotechnology* 28, 2014, pp. 127-133.

[8] D. Endy, "Foundations for engineering biology", *Nature* 438, 2005, pp. 449-453.

[9] D.G. Gibson, J.I. Glass, C. Lartigue et al., "Creation of a bacterial cell controlled by a chemically synthesized genome", *Science* 329, 2010, pp. 52-56.

[10] P.L. Luisi, F. Ferri, P. Stano, "Approaches to semi-synthetic minimal cells: a review", *Naturwissenschaften* 93, 2006, pp. 1-13.

[11] E. Altamura, P. Stano, P. Walde, F. Mavelli, "Giant vesicles as micro-sized enzymatic reactors: perspectives and recent experimental advancements", *International Journal of Unconventional Computing* 11, 2015, pp. 5-21.

[12] E.R. Kandel, Y. Dudai, M.R. Myford, "The molecular and Systems Biology of Memory", *Cell* 157, 2014, pp. 163-186.

[13] G. Seelig, D. Soloveichik, D.Y. Zhang, E. Winfree, "Enzyme-Free Nucleic Acid Logic Circuits", *Science* 314, 2006, pp. 1584-1588.

[14] E. Winfree. *Algorithmic self-assembly of DNA*. PhD thesis, California Institute of Technology, May 1998.

[15] E. Winfree, X. Yang, N.C. Seeman. "Universal computation via self-assembly of DNA: Some theory and experiments", *DNA Based Computers Two* 44, 1999, pp. 191-213.

[16] D.T. Gillespie, "Exact Stochastic Simulation of Coupled Chemical Reactions". *The Journal of Physical Chemistry* 81(25), 1977, pp. 2340-2361.

[17] A. Kocabas, C. Shen, Z.V. Guo, S. Ramanathan, "Controlling interneuron activity in Caenorhabditis elegans to evoke chemotactic behaviour", *Nature* 490, 2012, pp. 273-277.

[18] D. Jaeger, R. Jung (eds.), *Encyclopedia of Computational Neuroscience*, 2015, Springer, pp. 1-3229.

[19] G.N. Reeke, R.R. Poznanski, K.A. Lindsay, J.R. Rosenberg, O. Sporns (eds.), *Modeling in the Neuroscience - from Biological Systems to Neuromimetic Robotics*, 2005, CRC Press, pp. 1-711.