INTERFACE FOCUS

Introduction

Computation by natural systems

Dominique Chu[1], Mikhail Prokopenko[2] and J. Christian J. Ray[3]

[1]School of Computing, University of Kent, Canterbury CT2 7NF, UK
[2]Centre for Complex Systems, Faculty of Engineering and IT, University of Sydney, Sydney, New South Wales 2006, Australia
[3]Center for Computational Biology, Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045, USA

DC, 0000-0002-3706-2905; MP, 0000-0002-4215-0344

Computation is a useful concept far beyond the disciplinary boundaries of computer science. Perhaps the most important class of natural computers can be found in biological systems that perform computation on multiple levels. From molecular and cellular information processing networks to ecologies, economies and brains, life computes. Despite ubiquitous agreement on this fact going back as far as von Neumann automata and McCulloch–Pitts neural nets, we so far lack principles to understand rigorously how computation is done in living, or active, matter. What is the ultimate nature of natural computation that has evolved, and how can we use these principles to engineer intelligent technologies and biological tissues?

1. Introduction

In March 2018, we held a Royal Society Theo Murphy workshop to bring together innovators in the nature of non-traditional computation, broadly interpreted and from all career stages, to inspire new directions towards this pressing gap. We had researchers who have a common interest, but who do not normally meet, with the goal of creating a common research agenda. We learned from experimental synthetic biologists what current technologies allow, from complex systems theorists what existing principles can tell us, from computer scientists and statisticians what can be learned from noisy information transfer, and from physicists what non-equilibrium principles may apply to strongly out-of-equilibrium complex matter that breeds computation.

This issue of *Interface Focus* represents the culmination of that meeting, reflecting four major themes that arose during the meeting: (i) non-traditional computing devices; (ii) neural networks and neuronal information processing; (iii) cellular and molecular biological information processing; and (iv) the physics of information in complex systems. In this issue, we aim to create a road towards a new synthesis of natural computing, connecting perspectives on computation ranging from thermodynamics to biology.

2. Non-traditional computing devices

Nicolau and colleagues [1] discuss a strategy for reducing the computation time of combinatorial problems using network-based computation. Often, they require a brute-force approach and are NP complete. Combinatorial computation problems are ubiquitous and important, including such tasks as circuit verification, protein folding, formal reasoning and network problems such as routing. By formulating NP-complete problems as graphs, Nicolau and colleagues can design microfluidic network structures through which active agents (e.g. bacteria or filamentous proteins) can stochastically explore. The structure of the network encodes the problem. They present strategies for solving a host of different basic combinatorial problems in this way, including the subset sum problem, the clique problem, the Steiner tree problem, the travelling salesman problem, maze solving and the satisfiability problem. The subset sum problem, for example, asks, for a given set of n integers $S =$

THE ROYAL SOCIETY
PUBLISHING

$\{s_1, \ldots, s_n\}$, if there is a subset whose elements sum to a target integer T. Many problems in computer science are isomorphic to these, so potential improvements in calculating them is of high interest. What is the advantage of network computing over traditional electronic computers? In essence, they trade speed at the per-operation level for parallelizability.

The study of Adamatzky [2] argues that fungi *Basidiomycetes* can be used as computing devices. Adamatzky describes an architecture of 'fungal computers', within which a mycelium provides a network of processors, while fungal fruit bodies comprise an input/output interface. The study then demonstrates that information within a fungal computer can be represented by spikes of electrical potential, so that the electrical activity can facilitate a computation via the electrical impulses propagating in, and modified by, the mycelium network. One of the promising applications is large-scale environmental monitoring based on mycelium networks deployed in soil and air.

3. Neural networks and neuronal information processing

In their work, Saglietti et al. [3] bridge the fields of statistical mechanics, neuroscience and machine learning. One of the open problems in neuroscience is to understand how the brain can learn. One of the dominating models of learning goes back to J. J. Hopfield's influential paper in 1982 in which he showed how to encode memories in a simple disordered system. There are a number of shortcomings with Hopfield's network, in that it becomes unstable, shows spurious memories and can undergo catastrophic 'forgetting'. Saglietti et al. present a learning rule that overcomes many of these problems, thus providing novel insights into how the brain may learn, but also making a contribution to machine learning.

4. Cellular and molecular biological information processing

Suderman & Deeds [4] apply information theory to biochemical signal transduction networks. Processing information about their environment or even their internal state is an essential task of biological systems. Information transmission in cells, however, is noisy, which puts a limit on the amount of reliable information that can be processed. Indeed, measurements of the information that is transmitted by biological systems indicate that this is often less than a bit. This raises the question of whether this is a fundamental limitation of the networks. Suderman and Deeds use extensive simulations of artificial signalling networks to demonstrate that biological information processors transmit much less information than would be theoretically possible. This leaves open the question as to why this is the case. The authors suggest that extrinsic noise may reduce the information transmission capabilities of cells, and they even suggest that in certain situations lower information transmission is better than higher information transmission.

Schmelling & Axmann [5] explore a basic mechanism that cells have evolved to compute environmental shifts—the circadian rhythm. They argue that computational models have played an indispensable role in identifying unifying principles of circadian clocks. The central common feature is regulatory networks with multiple feedback loops, either with a post-translational oscillator as in cyanobacteria or with nested transcription–translation feedback loops as in eukaryotes. The cyanobacterial oscillator is the most primitive that is known and provides insights into how such a computational mechanism may have evolved. The oscillator consists of a monohexameric protein, KaiC, that gradually becomes phosphorylated during the day until it is completely phosphorylated at the onset of night. The reverse process occurs during the night until it is completely unphosphorylated in the morning. The daytime configuration stimulates the use of sunlight-derived energy for cell growth and global gene regulation. Schmelling and Axmann point out that environmental noise and seasonal variation are necessary for the evolution of circadian oscillators, and that oscillators provide a robustness mechanism to unpredictable changes in the environment. Circadian oscillators are a fascinating example of computation in nature that could inspire synthetic biology and non-traditional computing paradigms.

Wiesner and colleagues [6] address an ongoing problem of how developing cells in multicellular organisms compute the outcomes of their developmental programmes—how does a stem cell decide to proceed down, say, the erythroid path (committing to becoming B cells, T cells or natural killer cells, for example) as opposed to the myeloid path (committing to becoming red blood cells or macrophages, for example). A question of cellular identity arises: how much 'potential' for becoming different cell types does each step on the developmental pathway have? The best way to measure this may be using information theory—Shannon entropy tells of the level of disorder in the system. In a binary digital dataset, entropy is maximal when half the bits are 0 (or OFF) and half are 1 (or ON). Based on the high potential that stem cells have, intuition may suggest that cells lose entropy as they become more specifically ordered to a definite cell type. On the other hand, maturation of cell types may result in more genes being expressed. An increased system size may cause some maturation steps to increase rather than decrease the entropy. Wiesner and colleagues used single-cell transcriptomic datasets to measure the binary entropy of cells moving down developmental pathways. They found that the entropy in these datasets varies non-monotonically, increasing immediately after lineage commitment points. This suggests a new view of gene expression space during development, where new dimensions of expressivity open even as previous potential dimensions become inaccessible.

5. The physics of information in complex systems

Chu & Spinney [7] consider a physically plausible model of a so-called finite state machine or FSM. FSMs are a non-universal model of computation that has significant importance in computer science. The article discusses a possible physical implementation of FSMs and calculates the energy required to update such an FSM.

The key notion considered by Kolchinsky & Wolpert [8] is semantic information. The new formalization begins with the syntactic information captured by Shannon information which quantifies the amount of statistical correlation between

systems. It then defines the semantic information as the syntactic information between a physical system and its environment that is causally necessary for the system in order to maintain its existence over time. Semantic information is asymmetric: an organism may have semantic information about its environment, being dependent on its specific features, while the environment does not necessarily possess semantic information about the organism. The semantic information is further analysed from a thermodynamic point of view, using the methods of non-equilibrium statistical physics. In particular, one may define the semantic efficiency as the ratio of the semantic information observed under an optimal intervention to the system dynamics, to the total transfer entropy.

Harding *et al.* [9] study thermodynamics of contagions, considered as distributed computational processes. This perspective formalizes a disease spread developing across a contact network in statistical–mechanical terms, and identifies critical thresholds and distinct phases of epidemics. Analogously to the efficiency of heat-engine or refrigeration cycles, defined as the ratio of desired output (for example, the cooling effect) to the required work input, the study introduces the thermodynamic efficiency of contagions. This thermodynamic efficiency is defined as the ratio of the uncertainty reduction during an epidemic to the work needed by a specific intervention. This view is contrasted with an alternative interpretation according to which the thermodynamic efficiency of a pathogen emergence can be defined with respect to the work extracted by the pathogen during the infection spread. The study argues that knowing the efficiency of contagions would help in a comparative analysis of various interventions as well as pathogen emergence paths.

6. Outlook

We conclude with a discussion of the outlook for this field. The ever-increasing combined technological advances in machine learning and biotechnology raise some fascinating prospects: with high dimensional, high resolution data, we stand the chance of capturing what is 'really' happening in complex systems using machine-guided analysis pipelines. We also feel that the most important conceptual underpinning of biological science, evolution by natural selection, has been rather neglected in studies of biological computation. Nevertheless it lurks in the background and is the driving force for how natural computations arise in biology. A complete synthesis of biological computation must account for the evolutionary forces that shaped it.

Ongoing efforts to nucleate a community around the topic of computations in natural systems and non-traditional contexts have resulted in a new book [10] in addition to a wiki (https://centre.santafe.edu/thermocomp/Santa_Fe_Institute_Collaboration_Platform:Thermodynamics_of_Computation_Wiki). We hope that this meeting provides further inspiration for new directions in discovering how matter computes.

References

1. van Delft FCMJM, Ipolitti G, Nicolau Jr DV, Sudalaiyadum Perumal A, Kašpar O, Kheireddine S, Wachsmann-Hogiu S, Nicolau DV. 2018 Something has to give: scaling combinatorial computing by biological agents exploring physical networks encoding NP-complete problems. *Interface Focus* **8**, 20180034. (doi:10.1098/rsfs.2018.0034)

2. Adamatzky A. 2018 Towards fungal computer. *Interface Focus* **8**, 20180029. (doi:10.1098/rsfs.2018.0029)

3. Saglietti L, Gerace F, Ingrosso A, Baldassi C, Zecchina R. 2018 From statistical inference to a differential learning rule for stochastic neural networks. *Interface Focus* **8**, 20180033. (doi:10.1098/rsfs.2018.0033)

4. Suderman R, Deeds EJ. 2018 Intrinsic limits of information transmission in biochemical signalling motifs. *Interface Focus* **8**, 20180039. (doi:10.1098/rsfs.2018.0039)

5. Schmelling NM, Axmann IM. 2018 Computational modelling unravels the precise clockwork of cyanobacteria. *Interface Focus* **8**, 20180038. (doi:10.1098/rsfs.2018.0038)

6. Wiesner K, Teles J, Hartnor M, Peterson C. 2018 Haematopoietic stem cells: entropic landscapes of differentiation. *Interface Focus* **8**, 20180040. (doi:10.1098/rsfs.2018.0040)

7. Chu D, Spinney RE. 2018 A thermodynamically consistent model of finite-state machines.

Interface Focus **8**, 20180037. (doi:10.1098/rsfs.2018.0037)

8. Kolchinsky A, Wolpert DH. 2018 Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface Focus* **8**, 20180041. (doi:10.1098/rsfs.2018.0041)

9. Harding N, Nigmatullin R, Prokopenko M. 2018 Thermodynamic efficiency of contagions: a statistical mechanical analysis of the SIS epidemic model. *Interface Focus* **8**, 20180036. (doi:10.1098/rsfs.2018.0036)

10. Wolpert D, Kempes C, Grochow J, Stadler P (eds). 2018 *The interplay of thermodynamics and computation in natural and artificial systems*. Santa Fe, NM: Santa Fe Institute Press.