Preliminary Steps toward Artificial Protocell Computation

Barry McMullin Ciarán Kelly Darragh O'Brien George G. Mitchell
James Decraene

ALife Lab, RINCE, Dublin City University, Ireland
http://www.eeng.dcu.ie/~alife

Extended abstract submission for the International conference on Morphological Computation, 27-28 March 2007, ECLT, Venice Italy.

1 Introduction

Protocells are hypothesised as a transitional phase in the origin of life, prior to the evolution of fully functional prokaryotic cells [1, 2]. The work reported here is being done in the context of the PACE[1] project, which is investigating the fabrication of artificial protocells *de novo*. We consider here the important open question of whether or how articifial protocells (if or when they are successfully fabricated) might be applied as "computing" devices—what sort of computing might they be suitable for, and how might they be "programmed"? We also present some preliminary analysis of a crude model of such "evolutionary protocell computation".

2 Two Approaches to Protocell Computation

General purpose computational devices (hardware or software) are normally indefinitely large composites of components drawn from a relatively small and fixed repertoire (logic gates, instructions etc.). It follows that we can conceive protocell computation in at least two distinct ways:

- A protocell is considered as a "component"; suitably configured ("programmed") populations or colonies of protocells may then carry out diverse computations.
- A molecule is considered as a "component"; individual, suitably configured, protocells may then carry out diverse computations.

In this paper we concentrate exclusively on the second of these possibilities; with the further assumption that this cell-level computation might then be "programmed" via evolution among competing populations (lineages) of protocells.

3 What Kinds of Computation?

We can recognise certain forms of computation that have been successfully evolved by existing prokaryotes (the simplest free-living, single celled, organisms); specifically, the *cell-signalling networks* that function to control and regulate both the internal activity of the cell and its interaction with its environment. In the first instance then, we may hope to evolve protocells which incorporate this kind of real-time, signal processing, computation capability. Such capability would be of direct interest for at least some potential applications of protocells; and, in any case, would provide a basis for exploring more general computation in protocells.

[1] http://www.istpace.org/

4 MCS: A Minimal Model

We propose a highly simplified, minimal model, of protocell computation, loosely based on John Holland's Learning Classifier Systems [3, 4]). Classifier Systems can be viewed as string-rewriting systems. The operation of these systems depends on a population of strings or messages, and a rule-set which determines the action to take, given a particular string. In our Molecular Classifier System (MCS) we propose a similar string-rewriting paradigm of computation; but without the overt demarcation between "rules" and "messages". In MCS, any given, binary, string (which is taken to be analogous to an informational polymer molecule) may function either as a rule ("enzyme") or as a message ("substrate") according to different computational contexts.[2]

A protocell is then crudely modelled as a container for, among other things, a dynamic mix of essentially informational molecules, which continuously interact to generate new molecules (and which also decay or are actively destroyed). This "informational chemistry" may then be evolved to realise some particular computation—but only provided that it is simultaneously capable of sustaining its own dynamic organisation and of coupling actively and co-operatively with the other major protocell sub-systems (membrane and metabolism). In particular, this informational or computational sub-system must grow (in absolute number of molecules) and divide in sympathy with overall cell reproduction.

5 Initial Analysis

As an initial step in understanding the behaviour of such systems, we consider the case of informational molecules which are capable of (enzyme catalysed) replication (with a fixed, per-bit, error-rate). This allows the possibility of diverse, collectively autocatalytic [6], networks of molecules—such autocatalysis being an essential requirement both for the network to sustain itself against decay of individual molecules, and to grow in absolute molecular number to support cell reproduction. In the simplest case, we model the catalytic binding operation via a direct, binary, pattern match between strings, which effectively yields an indefinitely large family of molecular species which are individually autocatalytic (while still having a rich pattern of cross-catalytic interactions also).

In this "toy" case, we might immediately expect that the core, generic, behaviour of such an informational sub-system will be to exhibit a selectional, or "replicator", dynamics; and that any "network" would actually quickly collapse onto a single, dominant, "master" sequence, perhaps with a penumbra of other sequences representing the "quasi-species" in the immediate mutational neighbourhood of the master sequence [7]. This is the "selfish replicator" scenario in its most pure form [8]. Unfortunately, such degenerate, homogeneous, "networks" would clearly have minimal scope to realise interesting computation (which must surely involve a diversity of molecular species in more or less complex interaction with each other).

However, recall that this simple analysis neglects the fact that (unlike in the RNA world scenario) the informational sub-systems here are each coupled with, and completely dependant upon, the overall behaviour of their containing protocells. This means that viability and evolution at the protocell level should actually impose constraints on the dynamics of the informational subsystem within each cell. We can hypothesise that, in principle at least, this cell-level evolution may provide suitable conditions for the emergence and maintenance of *co-operation* in the informational sub-systems within each individual cell (in the manner characteristic of a "major evolutionary transition" [1]); which is to say, that networks of diverse molecular species, which realise significant and interesting computational behaviour, may be evolved and stabilised in this way.

The immediate goal of our work is therefore precisely to test this hypothesis in model systems.

This work is still at an early stage, and we have no substantive prognosis of the likely ultimate outcome to report as yet. But the first studies have shown than even the "simplest", single-cell (non-evolutionary) model exhibits somewhat more complicated phenomenology than had been

[2] cf. the "RNA world" hypothesis which proposes an epoch in the origin of life when a single family of polymer molecule—RNA or similar—could function both as an information carrier and as an active catalyst for transformation ("ribozyme") according to chemical context [5].

anticipated just on the basis of a pure replicator dynamics. In particular, even in the case of a molecular species that can bind to instances of itself, and thus catalyse its own ("self"-)replication, this is *not* equivalent to the case where a *single individual molecule* can catalyse its own replication (what we will here call "*auto*-replication"—cf. protein "auto-processing" [9]). This is a subtle, but significant distinction. It introduces an important density-dependent selection dynamic: in the case where pairwise interaction is required to support replication, then if any species transiently achieves even a slightly higher than average concentration, this can be quickly amplified and diversity can collapse, *even in the absolute absence of any "intrinsic" fitness differences between molecular species.* This effect is strong enough to dominate the dynamics even at relatively high replication error rates.

6 Conclusion

The fabrication of artificial protocells, if it can be achieved, has the potential to provide an entirely novel, nano-scale, substrate for the realisation of powerful new computational systems. However, the very novelty of this platform presents very complex, and as yet poorly understood, challenges to identify appropriate and effective ways of applying this new potential. The current work is an early, speculative, attempt to address some of these open questions through investigating highly simplified models of building and "programming" such "Artificial Protocell Computation". It is not expected to translate directly or immediately to chemical realisation; but it may still serve usefully to guide or constrain such eventual realisation.

Acknowledgements

This work has received financial support under EU FP Integrated Project PACE (contract number 002035).

References

[1] J. Maynard Smith and E. Szathmáry, *The Major Transitions in Evolution*. Oxford Press, 1997.

[2] S. Rasmussen, L. Chen, D. Deamer, D. Krakauer, N. Packard, P. Stadler, and M. Bedau, "Transitions from Nonliving to Living Matter," *Science*, vol. 303, no. 5660, pp. 963–965, 2004.

[3] J. Holland, *Adaptation in natural and artificial systems*. MIT Press Cambridge, MA, USA, 1992.

[4] J. Holland, "Exploring the evolution of complexity in signaling networks," *Complexity*, vol. 7, no. 2, pp. 34–45, 2001.

[5] G. Joyce, "The rise and fall of the RNA world.," *The New Biologist*, vol. 3, no. 4, pp. 399–407, 1991.

[6] B. McMullin, "Some remarks on autocatalysis and autopoiesis," *Annals of the New York Academy of Sciences*, vol. 901, pp. 163–174, 2000.

[7] M. Eigen, "Selforganization of Matter and the Evolution of Biological Macromolecules," *Die Naturwissenschaften*, vol. 58, no. 10, pp. 465–523, 1971.

[8] R. Dawkins, *The Selfish Gene*. Oxford: Oxford University Press, 1976.

[9] H. Paulus, "Protein splicing and related forms of protein autoprocessing," *Annual Review of Biochemistry*, vol. 69, pp. 447–496, 2000.