Evolution of Self-Maintaining Cellular Information Processing Networks

Barry McMullin[1] and James Decraene[2]

[1]The Rince Institute, Dublin City University, Ireland
[2]School of Computer Engineering, Nanyang Technological University, Singapore
barry.mcmullin@dcu.ie

Background

This contribution is an extended abstract of (Decraene and McMullin, 2011). What we here term Cellular Information Processing Networks (CIPNs) are biochemical systems of interacting molecules occurring in living cells. CIPNs are responsible for coordinating cellular activities in response to internal and external stimuli (e.g., chemotaxis signalling pathways). CIPNs can be regarded as special purpose computers (Bray, 1995). A single enzyme molecule effectively carries out pattern matching to identify and bind target substrate(s), and then executes a discrete computational operation in transforming these into the product molecule(s). The concept of *collective autocatalysis*, formulated by Farmer et al. (1986), denotes a collection of molecular species where each of them is the product of at least one reaction catalysed by at least one other species of the set. Fontana and Buss (1994) developed this into a more general formal concept of (collective) *self-maintenance*, and it has more recently been elaborated and refined in the *Chemical Organization Theory* of Dittrich and Speroni (2007). Self-maintenance ensures that reaction networks can reconstitute themselves when subjected to perturbations and during cellular divisions. It may thus mediate between the conflicting objectives of robustness and evolvability in reaction networks.

In contrast to modern living cells, the cellular model considered here does *not* incorporate a distinct genetic translation system. It is motivated by the (presumed) evolution of information processing in (proto-)cells *prior* to the emergence of the genetic architecture.

The Artificial Chemistry (MCS.bl)

We employ an agent/string-based Artificial Chemistry called the Molecular Classifier System (MCS.bl[1]) which is based on Holland's broadcast language (Holland, 1992, pp. 143-152). The basic elements (the abstract "molecules") are formally strings on a specified symbol alphabet ("atomic" species). Chemical reactions are stochastic (molecular "mutation" may alter the generated product strings), reflexive (no distinction made between operands and operators) and catalytic. Any single molecule may contain several condition/action rules which define its binding and enzymatic properties. In general the broadcast language allows arbitrary string transformations (computations) to be expressed; however, for the specific experiments described here, individual autocatalysis (self-catalysed replication) is explicitly disabled. Populations of molecules are encapsulated in containers to form "cells"[2]. Each cell functions as a separate well-stirred reactor. The number of molecules in a cell may increase until the cell matches a specified division criterion; a cell then divides with stochastic assortment of molecules into two daughter cells. Where particular molecular species are present in small numbers in a parent cell they may be absent completely in one of the two daughter cells, giving rise to distinct, cell-level, mutation events. The total number of cells is fixed: each division triggers the removal of another cell selected at random. The system is implemented on a small parallel computer cluster, with one CPU per cell. The real-time required for individual molecular interactions may vary with the specific detailed structures of the molecules involved. Cell reproduction rate is dependent on the real-time rate of catalytic reactions occurring in the cell, and on the specific criteria in effect for cell division. Distinct, interacting, selectional dynamics arise at both molecular and cellular levels.

Experiment: Molecular Amplification

In the first experiment cells are evolved to carry out amplification of a given molecular species. This is motivated by conceptually similar *in vivo* investigations reported in the literature (Fong et al., 2005). The cell division criterion is configured so that cells divide when a target molecular species (s_T) reaches a threshold number of instances. The cellular reproduction rate (fitness) therefore depends on the ability of the cell to promote the growth of s_T while still preserving overall collective self-maintenance of all required molecular species. The system is initialised ("seeded") with

[1]MCS.bl source code and documentation is available at: http://esignet.net/dokumente/upload/WP13

[2]For brevity, we say simply "cell" here; but this should be read as "proto-cell" throughout.

a hand-designed, viable cellular species (self-maintaining at the molecular level, including the target molecular species, so that cellular division is possible).

Similar phenomena are encountered in multiple runs. One typical run is described and analysed in more detail. In the course of this run, 1235 different and unique cellular species were generated in total, but of these, just three successively came to dominate the cellular population, through three identifiable displacement events. Careful analysis of both the molecular dynamics within the dominant cellular species and the cell-level population dynamics allowed determination that the first observed displacement in this run was selectional, with a clear increase in fitness (molecular amplification function); but the subsequent two displacements represented drift among essentially equal-fitness cellular species. That said, more fine-grained examination of the displacement events also shows that they are correlated with significant transient increases in cellular species diversity. These displacements are thus significantly more intricate than straight selection or drift between two "pure" cell lines. In effect, a single molecular mutation event can give rise to a complex cascade of cell-level mutations.

Experiment: Crosstalk

Crosstalk phenomena arise very naturally in real biochemical information processing networks due to the fact that molecules from different signalling pathways may share the same physical reaction space (the cell). Depending on the relative specificities of the reactions there is then an automatic potential for any given molecular species to contribute to signal levels in multiple pathways. Here we describe an experiment investigating the evolutionary dynamics arising when distinct cells, with potential crosstalking pathways, are forcibly merged, but subsequent cell division is constrained to maintain selected molecular components from *both* pre-existing reaction networks (so cellular species in which one network simply displaces the other cannot continue to reproduce). This work is naturally related to the symbiogenesis theory which was originally postulated by Mereschkowsky (1910), and already explored computationally by Barricelli, on the first stored program digital computers, in the 1950's (Barricelli, 1957).

Over a number of runs, various common features are observed under these experimental conditions. A very rich variety of cellular species emerges, and in general there is significantly more cellular species diversity than in the previously described experiment: it is rare in this case for a single cellular species to exceed more than half of the population. Nonetheless, distinct displacement events can still be observed; and it is possible to analyse the molecular behaviour of a selection of mutant cellular species in detail. It is typical to observe the emergence of cellular species containing a "meta-reaction network", still including all the seed molecular species, but also additional molecular species, exploiting crosstalk and bridging between the seed species, and participating in the collective self-maintenance. In this sense, this experiment demonstrates a some (limited) evolutionary growth in the complexity of the self-maintaining reaction networks — both in terms of number of species and number of reactions composing the network. It is also observed that the gestation time of the dominant cellular species successively decreases

Conclusion

We have presented a preliminary investigation of the role of collective self-maintenance in the evolution of (proto-)cellular information processing reaction networks. To assist this research, we built a novel agent-based multi-level selectional Artificial Chemistry. This was applied to the evolution of a single and multiple/crosstalking self-maintaining reaction networks. In these experiments, cellular species were successfully evolved to achieve the pre-specified information processing functions more effectively and exhibited a relatively higher level of complexity (by at least some reasonable measures). This proof of concept should contribute, to some extent, to understanding of the much more general problem of *open-ended* evolutionary growth of complexity using Artificial Chemistries.

Acknowledgements

This work was supported by the ESIGNET Project (Evolving Cell Signalling Networks *in silico*), funded under the EU FP6 NEST initiative, contract no. 12789.

References

Barricelli, N. (1957). Symbiogenetic Evolution Processes Realized by Artificial Methods. *Methodos*, 9(35-36):143–182.

Bray, D. (1995). Protein Molecules as Computational Elements in Living Cells. *Nature*, 376(6538):307–312.

Decraene, J. and McMullin, B. (2011). The evolution of complexity in self-maintaining cellular information processing networks. *Advances in Complex Systems*, 14(1):55–75. Available from: `http://doras.dcu.ie/16292/`.

Dittrich, P. and Speroni, P. (2007). Chemical Organisation Theory. *Bulletin of Mathematical Biology*, 69(4):1199–1231.

Farmer, J., Kauffman, S., and Packard, N. (1986). Autocatalytic Replication of Polymers. *Physica D*, 22(1-3):50–67.

Fong, S., Burgard, A., Herring, C., Knight, E., Blattner, F., Maranas, C., and Palsson, B. (2005). In Silico Design and Adaptive Evolution of Escherichia coli for Production of Lactic Acid. *Biotechnology and bioengineering*, 91(5):643–648.

Fontana, W. and Buss, L. (1994). "The Arrival of the Fittest": Toward a Theory of Biological Organization. *Bulletin of Mathematical Biology*, 56(1):1–64.

Holland, J. (1992). *Adaptation in Natural and Artificial Systems*. MIT Press Cambridge, MA, USA, 2nd edition. See also Holland(1975).

Mereschkowsky, C. (1910). Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. *Biol. Centralbl*, 30:278–288.