Exploring evolutionary stability in a concurrent artificial chemistry

James Decraene, George Mitchell, Barry McMullin

Artificial Life Laboratory - Research Institute for Networks and Communications Engineering
School of Electronic Engineering - Dublin City University, Ireland
Email: james.decraene@eeng.dcu.ie

Introduction

Multi-level selection has proven to be an affective mean to provide resistance against parasites for catalytic networks (Cronhjort and Blomberg, 1997). One way to implement these multi-level systems is to group molecules into several distinct compartments (cells) which are capable of cellular division (where an offspring cell replaces another cell). In such systems parasitized cells decay and are ultimately displaced by neighboring healthy cells. However in relatively small cellular populations, it is also possible that infected cells may rapidly spread parasites throughout the entire cellular population. In which case, group selection may fail to provide resistance to parasites. In this paper, we propose a concurrent artificial chemistry (AC) which has been implemented on a cluster of computers where each cell is running on a single CPU. This multi-level selectional artificial chemistry called the Molecular Classifier Systems (Decraene et al., 2007) was based on the Holland broadcast language (Holland, 1992). An attribute inherent to such a concurrent system is that the computational complexity of the molecular species contained in a reactor may now affect the fitness of the cell. This molecular computational cost may be regarded as the chemical *activation energy* necessary for a reaction to occur. Such a property is often not considered in typical Artificial Life models. Our experimental results obtained with this system suggest that this activation energy property may improve the resistance to parasites for catalytic networks. This work highlights some of the benefits that could be obtained using a concurrent architecture on top of computational efficiency. We first briefly present the Molecular Classifier Systems, this is then followed by a description of the multi-level concurrent model. Finally we discuss the benefits of using this multi-level concurrent model to enhance evolutionary stability for catalytic networks in our AC.

Molecular Classifier Systems

Molecular Classifier Systems are a class of string-rewriting based AC (inspired by Learning Classifier Systems-LCS) that were proposed to investigate protocell computation (McMullin et al., 2007). As opposed to traditional string-rewriting systems, operations are stochastic and reflexive (no distinction made between operands and operators). The behavior of the condition (binding) properties and action events (enzymatic functions) is defined by a language specified within the MCS. This "chemical" language defines and constrains the complexity of the chemical reactions that may be modeled and simulated. In this AC, all reactants are catalytic in the sense that they are not consumed during reactions. These reactions result from successful molecular interactions which occur at random. When a reaction occurs, a product molecule is inserted into the reactor, however if the reactor is saturated then another molecule (selected at random) may be removed from the reactor space (designating the system outflow).

A molecule may contain several condition/action rules which define the binding and enzymatic properties. A reaction between molecules occurs if at least one conditional part from any rules in a molecule A matches a target molecule B. A is regarded as an enzyme whereas B is regarded as a substrate molecule. When a reaction occurs, the action part from the satisfied rule in A is utilized to perform the enzymatic operations upon the bound substrate molecule B. This operation results in the production of another offspring (product). If several rules in A are satisfied by B, then one of these rules is picked at random and employed to carry out the enzymatic function. The computational complexity of a given molecule depends on its string length and nature of different operands.

A detailed description is purposely omitted in this paper as it may confuse the present issue. Our intention is to describe the potential benefits of using an overall (concurrent) system architecture instead of focusing on specific AC implementation details.

Multi-level selectional and concurrent architecture

In a single-level selectional model, all molecules are contained in a single reactor in which they were competing with each other. This molecular competition is referred to as the

first level of selection. In the multi-level selectional model, we introduce multiple reactors where each of them contains a population of molecules. These reactors or cells may be subjected to cellular division, which results in the replacement of the parent reactor and creation of two offspring cells. However, the number of reactors in the universe is fixed. As a result such a cellular division also triggers the removal of another reactor selected at random. In a similar manner to molecules, cells are competing with each other which is regarded as the second level of selection.

In contrast to the single level model, successful reactions do not lead to the removal of a random molecule in the reaction space. Thus the number of molecules contained in a reactor may increase until it reaches a finite limit l. When a reactor is saturated, a division occurs as follows: Half of the molecules contained in the reactor are selected at random, then these molecules are removed from this reactor and are inserted into the offspring cell. This newly created cell is then inserted into the population of reactors. Finally, a reactor is picked at random (other than the offspring and parent cell) and removed from the reactor population.

Furthermore this multi-level model was implemented as a concurrent system where each reactor is run on a single CPU. In this concurrent model, the fittest reactors would not only be the reactors that exhibit a high molecular growth rate, but reactors that also contain molecules that are fast to compute. In other words, if we consider two reactors which present an equal overall molecular growth rate, but contains molecules with different computational complexities, the reactor which possesses a smaller overall molecular computational complexity will have the selective advantage. This computational complexity may thus be regarded as the chemical activation energy. The higher computational cost is, the more energy is necessary for a given chemical reaction to occur.

Activation energy enhancing evolutionary stability at the cellular level

Previous results (Decraene et al., 2008) indicated that in simple evolutionary simulations of the single level Molecular Classifier System, parasites would rapidly emerge and lead to a degenerated scenario: We observed the emergence of fitter mutants that were able to parasite the master molecule. This increase in molecular fitness was associated with an increase in the molecular string length. However successive emergences of fitter molecules lead to ever longer molecules which has an effect of increasing the mutation rates. As the mutation rate increases, disruptive structural changes occur more often and may "break" the molecules function. This *elongation catastrophe* phenomenon ultimately leads to the system level extinction. To illustrate this case, let us consider the following possible reactions which lead to such an elongation catastrophe:

$$\begin{aligned} R + R &\longrightarrow 3R \\ R + R' &\longrightarrow R + 2R' \\ R' + R &\longrightarrow R + 2R' \\ R' + R' &\longrightarrow 2R' + R'' \\ &\cdots \end{aligned}$$

R' is a mutant of R, R' possesses a selective advantage over R as described in above reactions. R' has a longer molecular string than R so has R'' over R'. Similarly R'' has a selective advantage over R and R'. Two molecules R'' reacting with other would lead to the production of a molecule R''', which shows the progressive increase in length of the molecular species.

To best exemplify the benefits of using a concurrent AC for a given multi-level selectional model, we propose to consider two simple experimental cases where only two compartments are employed. In both cases only R molecules are seeded in a reactor A whereas R' molecules are inserted in a reactor B with equal initial concentrations. Reactor A is therefore considered as the "healthy" cell whereas reactor B the "infected" cell. Both reactors have the same maximum carrying capacity.

- In the first case, the multi-level selectional model is implemented on a single CPU system where no selective advantage is given to molecules contained in reactor A or B. In both reactors, any interactions lead to a successful reaction as any molecules may react with any other molecules contained in a given reactor. Therefore both reactors have the same molecular growth and thus equal fitness. Over time, when one of the two reactors reaches the division threshold (i.e., saturated), it then triggers its division which leads to the displacement of the other cell. Within these conditions, over a series of experiments, we may confidently say that in half of the cases, reactor A would first displace reactor B and ensure the survival of replicases R. However in the other cases, reactor B would have displaced reactor A, resulting in population of mutant molecules subjected to the elongation catastrophe phenomenon.

- In the second case, this multi-level selectional model is implemented as a concurrent system where reactor A and reactor B are running on two distinct CPUs. No selective advantage is given to either reactor A or reactor B (both CPUs running on exact same computing infrastructures). In this case, reactor B would rapidly be filled with mutant molecules having exponentially increasing string length. The reactor B would require then more computational time or activation energy to reach the division threshold than reactor A. As a result, reactor B would always displace reactor A.

In the first scenario, we noted a high probability that infected cell may displace the healthy reactor. Whereas with the concurrent system, due to the activation energy property affecting the molecules and reactors fitness, the healthy reactor would always displace the infected reactor. As a result, thanks to a activation energy property given for free with the concurrent architecture, the evolutionary stability of our system was enhanced.

In this simple case, only two reactors were considered to facilitate the understanding of the benefits of our concurrent approach. Of course, if the cellular population was sufficiently larger then group selection only would have provided an efficient resistance to parasites. However in relatively small cellular populations (which could be envisaged in prebiotic conditions), parasitized cell may still be able to infect the entire population. As a result, if we consider the conditions we described above, group selection may fail and degenerated outcomes may occur preventing stable cooperation between molecular species.

Moreover in traditional multi-level model infected cells would usually be displaced only when these have decayed (exhibiting a quasi null overall reaction rate). Whereas in this multi-level concurrent AC, infected cells (having a higher energy demand) would rapidly be displaced by healthy cells. Making this overall molecular system more reactive against parasitism.

Conclusion

We presented a multi-level concurrent Artificial Chemistry based on Molecular Classifier Systems and Holland broadcast language. We demonstrated how such a concurrent architecture may provide am activation energy property for free. It was then shown that this property provided the benefit of enhancing evolutionary stability for catalytic networks in our system. Future work includes the introduction of a single level selectional system where only compartmentalization and molecular diffusion between compartments are considered. It was shown (McCaskill et al., 2001), through the use of an analytical model, that such an approach also enhances evolutionary stability. Comparisons between our multi-level concurrent system and this spatially resolved molecular system will follow. Hybrid systems combining both approaches will also be examined.

Acknowledgment

This work was funded by ESIGNET (Evolving Cell Signaling Networks in Silico), an European Integrated Project in the EU FP6 NEST Initiative (contract no. 12789). The experiments described in this paper were performed on computing facilities provided by the ESIGNET project and by the Irish Center for High-End Computing (ICHEC, project code: dccom011c). We are also grateful to fruitful discussions and comments from Ciaràn Kelly.

References

Cronhjort, M. and Blomberg, C. (1997). Cluster compartmentalization may provide resistance to parasites for catalytic networks. *Physica D: Nonlinear Phenomena*, 101(3-4):289–298.

Decraene, J., Mitchell, G. G., and McMullin, B. (2007). Evolving artificial cell signaling networks: Perspectives and methods. In Dressler, F. and Carreras, I., editors, *Advances in Biologically Inspired Information Systems*, volume 69 of *Studies in Computational Intelligence*, pages 165–184. Springer.

Decraene, J., Mitchell, G. G., and McMullin, B. (2008). Unexpected Evolutionary Dynamics in a String-Based Artificial Chemistry. *Artificial Life XI Proceedings of the eleventh International Conference on the Simulation and Synthesis of Living Systems, S. Bullock, J. Noble, R, Watson, M. Bedau (Eds.). MIT Press*. To appear.

Holland, J. (1992). *Adaptation in natural and artificial systems*. MIT Press Cambridge, MA, USA.

McCaskill, J., Füchslin, R., and Altmeyer, S. (2001). The Stochastic Evolution of Catalysts in Spatially Resolved Molecular Systems. *Biological Chemistry*, 382(9):1343–1363.

McMullin, B., Kelly, C., OBrien, D., Mitchell, G. G., and Decraene, J. (2007). Preliminary Steps toward Artificial Protocell Computation. In *MorphComp 2007*, Venice, Italy.