CLOSURE IN ARTIFICIAL CELL SIGNALLING NETWORKS Investigating the Emergence of Cognition in Collectively Autocatalytic Reaction Networks

James Decraene

Artificial Life Laboratory, Research Institute for Networks and Communications Engineering Dublin City University, Glasnevin, Dublin, Ireland

Keywords: Cell Signalling Networks, Closure, Evolutionary Dynamics, Minimal Cognition.

Abstract: Cell Signalling Networks (CSNs) are complex biochemical networks responsible for the coordination of cellular activities in response to internal and external stimuli. We hypothesize that CSNs are subsets of collectively autocatalytic reaction networks. The signal processing or cognitive abilities of CSNs would originate from the closure properties of these systems. We investigate how Artificial CSNs, regarded as minimal cognitive systems, could emerge and evolve under this condition where closure may interact with evolution. To assist this research, we employ a multi-level concurrent Artificial Chemistry based on the Molecular Classifier Systems and the Holland broadcast language. A critical issue for the evolvability of such undirected and autonomous evolutionary systems is to identify the conditions that would ensure evolutionary stability. In this paper we present some key features of our system which permitted stable cooperation to occur between the different molecular species through evolution. Following this, we present an experiment in which we evolved a simple closed reaction network to accomplish a pre-specified task. In this experiment we show that the signal-processing ability (signal amplification) directly resulted from the evolved systems closure properties.

1 INTRODUCTION

Cell Signalling Networks (CSNs) are biochemical networks of interacting molecules (proteins, ions, secondary messengers, etc) occurring in living cells. Through complex molecular interactions (e.g., signal transduction), CSNs are able to coordinate critical cellular activities (e.g., cell differentiation, apoptosis) in response to internal and external stimuli.

In the presence of a genetic subsystem, the production of CSN components may result from the translation of the genetic code. However to coordinate this decoding process, a newborn cell still requires a minimal CSN which would be inherited from the parent cell during cellular division. In a prebiotic world, in which genetic material may potentially not be present, "primitive" CSNs would need an alternative mechanism to ensure their production and maintenance.

We may argue that both of these minimal/primitive CSNs from the post/prebiotic world may have to replicate themselves prior to the cellular division. This would allow the replicated CSNs to be "distributed" to the offspring cells. Errors may occur during the replication process, e.g., an offspring cell may inherit only a partial or mutant CSN. Thus resulting in potentially defective cells which would lead to a variety of undesired effects (e.g., premature cell death).

As a result, the "fitness" of a cell is implicitly represented by the *survival* and *performance* of a cell in achieving self-maintenance *and* cell-level replication. Based on the above assumption, we hypothesize that CSNs may be regarded as subsets of collectively autocatalytic (closed) reaction networks. In these closed and synergetic reaction networks, self-maintenance is achieved through the systems ability to produce the catalysts and substrates for all its reactions. This catalytic closure property originates from the autocatalytic set theory, which is one of the current candidate theories explaining the origins of life (Kauffman, 1997).

These closed reaction networks would have the additional ability to replicate themselves as a whole (during cellular division). Through the evolutionary process, we suggest that these systems may develop signal processing abilities which are typical of real CSNs. However these Artificial CSNs (ACSNs) must preserve closure to self-maintain and survive. Thus the systems signal-processing abilities may originate from the closure properties where inputs (internal or external stimuli) are perturbations to the system. The system outputs would result in a modification of the systems state and of its closure properties.

Following (van Duijn et al., 2006) we regard the coupling of these ACSNs signal-processing abilities and closure properties as minimal principles for life and cognition. By examining the conditions for the emergence of distinctive ACSNs signal-processing abilities, we investigate the emergence and origins of minimal cognition. These ACSNs may thus be regarded as Artificial Cognitive Systems which may potentially provide a range of applications in computing and engineering. Thus the goals of this research are twofold and can be summarized as follows:

- From a theoretical point of view: to study the origins and complexity evolution of cognition in biological systems.
- From a practical perspective: to provide a new computational paradigm inspired by nature.

To assist this research we developed the MCS.bl, a string-based Artificial Chemistry (AC) based on the Molecular Classifier Systems (MCS) (McMullin et al.,) and the Holland broadcast language (Holland, 1992). This technique examines the evolution of biochemical networks from an undirected, selfengineered and autonomous perspective. In this approach no explicit fitness function is defined, the evolutionary process is driven by the capacity of CSNs to self-maintain and perform the necessary signalprocessing operations or "computation" to achieve cell-level replication.

A key step of this research is to identity the conditions which provide evolutionary stability to the closed reaction networks when simulated in the MCS.bl. Examining such phenomena relates closely to other studies which have been conducted on Holland's Alpha-Universes (Holland, 1976), Tierra (Ray, 1991) and Alchemy (Fontana and Buss, 1994b). In this paper we first describe the MCS.bl. We then discuss how a multi-level and concurrent model enhanced the evolutionary stability of the system. Finally we present an experiment in which we evolved a minimal closed ACNS to accomplish a pre-specified task (signal amplification).

2 MOLECULAR CLASSIFIER SYSTEMS

Molecular Classifier Systems are a class of stringrewriting based AC inspired by Learning Classifier Systems (LCS). As opposed to traditional stringrewriting systems, operations are stochastic and reflexive (no distinction made between operands and operators). The behaviour 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 modelled 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.

We proposed a simplification of the Holland broadcast language (Decraene et al., 2007) which is used as the MCS chemical language resulting in the MCS.bl system. The MCS.bl differs from the original MCS (McMullin et al.,) by introducing more complex chemical reactions (only replications may occur in the MCS). 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.

A number of differences exist between our simplified broadcast language (BL) and the LCS, e.g., the LCS's alphabet is $\lambda = \{1, 0, \#\}$ whereas the BL includes additional symbols $\Lambda = \{1, 0, *, :$, \Diamond , \triangle , ', ∇ }. The basic elements of the BL are strings made from Λ called *broadcast devices*. A broadcast device is parsed into zero, one or more broadcast units, where each unit represents a single condition/action rule. The symbol * separates broadcast units within a broadcast device. The symbol : separates a condition from an action within a single broadcast unit. Os and 1s are basic informational symbols. $\{\Diamond, \nabla, \Delta\}$ are single/multiple character(s) wildcards that may also transpose matched strings into output strings. Quoted symbols (preceded by ') are prevented from interpretation. Figure 1 depicts an example broadcast device which may bind and react with a copy of itself, this reaction is presented in Figure 2.

A detailed description is omitted in this paper, see



Figure 1: An example broadcast device.



Figure 2: Example reaction.

Table 1: Examp	le operations	realized	with the	MCS.bl.
----------------	---------------	----------	----------	---------

Enzyme	substrate	product	operation
*⊽1: ⊽0	1:0	0	no reaction
⊽1:′▽	0:1	*0:1	activation
*′*0⊽:0⊽	*0:1	0:1	inhibition
$* \nabla : \nabla$	*00:11	*00:11	universal replication
$* \nabla 0 : \nabla 0$	*∇0:∇0	*∇0:∇0	self-replication
*⊽1:⊽10	*0:1	*0:10	concatenation
*∇1:∇	*0:1	*0:	cleavage

(J.Decraene, 2006) for full specification of our BL implementation. Table 1 presents a number of example operations that can be realized with the MCS.bl.

3 EXAMINING EVOLUTIONARY STABILITY

As this evolutionary system is an undirected approach, the first key step was to obtain catalytic networks that are able to self-sustain over time. In order to achieve such evolutionary robustness, it is necessary for these networks to possess mechanisms which provide some protection from parasites that may be formed. In this section we examine the system evolutionary stability firstly when implemented as a single level selectional model, secondly when the MCS.bl is implemented as a multi-level concurrent system.

3.1 Single-level Selectional Model

In a single-level selectional model, all molecules are contained in a single reactor in which they are competing with each other. This molecular competition is referred to as the first level of selection. Each time a successful reaction occurs, a product molecule is inserted into the reactor, if the latter is saturated then another molecule (selected at random) is removed from the reactor space (designating the system outflow), see Figure 3.1.

A formal explanation of this outcome is given by modelling the system with the (approximate, continuous) catalytic network equation (Stadler et al., 1993). The state of the system is described by the concentration vector $x = (x_1, ..., x_n)$ with $x_1 + ... + x_n = 1$ and $x_i > 0$, where x_i refers to the concentration of a molecular species (or collection of "chemically equivalent" species) s_i . The general dynamic behaviour is then given by:

$$\dot{x_k} = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij}^k x_i x_j - x_k \sum_{i,j,l=1}^{n} \alpha_{ij}^l x_i x_j$$
(1)
with $k = 1, \dots, n$

 α_{ij}^k are the rate constants for each reaction $s_i + s_j \rightarrow s_i + s_j + s_k$. In here, these simplify to:

$$\alpha_{ij}^{k} = \begin{cases} 1 \text{ if } s_i + s_j \to s_i + s_j + s_k \\ 0 \text{ otherwise} \end{cases}$$
(2)

Previous studies (Decraene et al., 2008) demonstrated that the spontaneous emergence and selfmaintenance of replicases in the MCS.bl was unlikely to occur. Consequently experiments were conducted in which, as in the Tierra system, a hand-designed molecule called the ancestor was introduced into the population. This replicase molecule was handdesigned so as to possess a high binding specificity, thus preventing exploitation from other molecules. This specificity ensures a rapid dominance over a randomly generated molecular population (which would not be possible if a universal replicase, having no binding specificity, was employed instead; see (Decraene et al., 2008)). Various molecular mutations through successful reactions were possible and a constraint over the molecular complexity was set due to computational limitations (experimental details and mathematical analysis may be found in (Decraene et al., 2008)). Although this MCS.bl implementation



Figure 3: Flowchart of the single-level selectional model.

was to some extent similar to the Tierra system, we did not observe the emergence of collectively autocatalytic networks. Instead we observed a degenerate outcome characterized by an "elongation catastrophe" phenomenon:

During evolution, a series of parasitic displacements was observed leading to ever longer molecules. However, as the mutation effect increases (with the length of a molecule), disruptive structural changes may also occur more frequently. We observed that consequences of these structural changes were twofold:

• Molecules may become inactive, thus being con-

sidered as substrate or waste molecules. These waste molecules would then accumulate and prevent any further enzymatic activities. A potential consequence of the BL syntax is that a mutation may disrupt dramatically the behaviour of an enzymatic molecule. This degenerative effect may be regarded as the consequence of the BL "brittleness" (asymmetric mutation effects (Ray, 1992)).

• The binding specificity may be increased. Mutations leading to the insertion of informational symbols such as 0s and 1s may also increase significantly the binding specificity of molecules. As a result, although these molecules may still possess an active site capable of some enzymatic function, they could not catalyse any molecules in the reactor.

Both of these phenomena result in a continual decrease in the overall reaction rate until reactions cease completely (i.e., the system decays and ultimately reaches a state of stasis). These results obtained with this single-level AC exhibited unexpected evolutionary dynamics which resulted in various degenerate cases: No stable cooperation between the molecular species could be observed in the evolutionary simulations. Thus this first implementation of our evolutionary system suffered from a critical lack of "robustness".

3.2 Multi-level Selectional and Concurrent Model

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 cell and creation of two offspring cells. However, the number of cells in the universe is fixed. As a result such a cellular division also triggers the removal of another cell 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 m of molecules contained in a cell may increase until the cell is full (i.e., when m is equal to the cell maximum capacity c). When a cell is full, a division occurs as follows: Half of the molecules contained in the cell are selected at random, then these molecules are removed from this cell and are inserted into the offspring cell. This newly created cell is then inserted into the cellular population. Finally, a cell is picked at



Figure 4: Flowchart of the multi-level and concurrent model.

random (other than the offspring and parent cell) and removed from the cell population, see Figure 3.2.

Furthermore this multi-level model was implemented as a concurrent system where each cell is run on a single CPU. In this concurrent model, the fittest cells would not only be the cells that exhibit a high rate of successful reactions (when compared to the total number of molecular collisions), but also cells that contain molecules that are fast to compute. For example let us consider two cells containing complete reaction networks (i.e., all molecular collisions lead to the successful production of molecules). Those cells would moreover contain molecules having different computational complexities. In here the cell which possesses a smaller overall molecular computational complexity will have the selective advantage. This computational complexity introduces in our model a notion of chemical kinetics and may alter the cellular growth rate (i.e., the cells fitness).s

We extended the MCS.bl with the multi-level concurrent model described above, our results indicated that none of the evolved cells resulting from the simulations suffered from elongation catastrophe. The nature of the evolved populations resulting from the simulation runs were equivalent to those expected from systems such as Alchemy: We observed the rapid domination of molecular organizations which involved a limited number of replicases capable to self-sustain over time.

Moreover, in traditional multi-level selectional models infected cells would usually be displaced only when these have decayed (Cronhjort and Blomberg, 1997; Hogeweg and Takeuchi, 2003). Whereas in our multi-level concurrent AC, infected cells (having a smaller growth rate) would rapidly be displaced by healthier cells (having a higher growth rate) due to the concurrent nature of the system. This concurrency property increased the system's ability to control parasitism.

4 EVOLVING CLOSED REACTION NETWORKS

The previous examination on evolutionary stability provided us with the necessary understanding to evolve closed reaction networks that are able to selfsustain over time. These closed reaction networks are regarded as primitive ACNS. Preliminary experiments suggested that when self-replication is enabled, evolved networks would present only limited interest. As these networks would converge toward molecular organizations involving a small number of altruist replicases. In the following experiment we disable self-replication, and employ a minimalist (handdesigned) collectively autocatalytic reaction network, see Figure 5.A. Let $S1 = *\nabla 0 : \nabla 1$, $S2 = *\nabla 0 : \nabla 0$, $S3 = *\nabla 1 : \nabla 1$, $S4 = *\nabla 1 : \nabla 1$, $S5 = *\nabla \diamond : \nabla 1$ and $S6 = *\nabla \diamond : \nabla 0$.

In section 3.2, a cell would divide only when sat-



Figure 5: A: Seed closed reaction network. B: Evolved closed reaction network promoting growth of molecular species *S*1.



Figure 6: Dynamics of seed closed reaction network, the species *S*1 and *S*4 overlapping (top line) and *S*2 and *S*3 also overlapping.

urated. We define a new cellular division criterion: Here a cell would divide when the amount of molecular species S1 reaches 200. The cellular growth rate now depends on the molecular growth rate of S1. The ability of the closed reaction network to promote the growth of S1 defines the cells fitness (i.e., a cell would increase its fitness by exhibiting a higher S1 growth rate). Thus the pre-specified task assigned to these simple closed ACSNs is to amplify the "signal" S1. We may also regard S1 as a necessary molecule such as a membrane molecule to allow the cellular division to occur. We first present the dynamics of the seed closed reaction network when the initial amount of each molecular species is equal to 10, see Figure 6. This graph was obtained by solving the ODE system generated by the reaction network (in SBML format) using the SBML ODE solver (Machné et al., 2006).

An evolutionary simulation was conducted in which 31 concurrent cells were employed. These



Figure 7: Dynamics of evolved closed reaction network, the species *S*5 and *S*6 overlapping (middle line), *S*1 is the top line and *S*4 is the bottom line. Each molecular species amount is initialized to 10.

cells were seeded with the molecular species (with initial amount of molecules 1) depicted in Figure 5.A. This experiment was run for one hour, on average the number of interactions per cell per hour was over 4.10^6 . The resulting dominant molecular organization is depicted in Figure 5.B. We first observed that this evolved network successfully maintained closure. We note that both species *S*2 and *S*3 have been replaced by new molecular species *S*5 and *S*6. These new molecular species increased the overall number of reactions from 6 to 9. We compare the dynamics of this evolved closed reaction network against the seed closed reaction network, see Figure 7.

In Figure 7, it can be seen that *S*1 reaches the division threshold at $t \approx 0.0422$ whereas in the seed closed reaction network *S*1 would reach this threshold at $t \approx 0.0588$. By reaching this threshold at an earlier stage, the evolved network gained a selective advantage over the seed network. We observed the emergence of molecular species *S*5 and *S*6 which had for effect to promote the growth of *S*1 whilst maintaining closure. The network closure properties evolved and allowed the network to promote the growth of species *S*1.

Moreover results indicated that when a cellular division occurs, the amount of the different molecular species would be optimized, see Figure 8. This optimization resulted in the ability to encourage the growth of *S*1 while limiting the growth of species *S*4, *S*5 and *S*6. However this amount distribution would not lead to the displacement of *S*4, *S*5 and *S*6 as these species are necessary to maintain closure.

In Figure 8, it can be observed that S1 would reach the division threshold when $t \approx 0.00973$. This optimized amount distribution provided this network a selective advantage over networks having differing



Figure 8: Typical dynamics of an evolved closed reaction network with optimized initial molecular amounts resulting from a cellular division



Figure 9: Comparison of S1 dynamics with differing networks and initial molecular amount distributions. $S1_a$ and $S1_b$ depicts the dynamics of S1 with initial molecular 10, using the seed reaction network and evolved reaction respectively. $S1_c$ represents the dynamics of S1 using the evolved reaction network and optimal initial molecular amount distribution.

amount distributions (as shown in Figure 7).

Finally Figure 9 compares the different dynamics of *S*1 using the different networks and initial number of molecules. This comparison highlights the successive improvements in the networks fitness (i.e., hastened the production of molecules A).

5 DISCUSSION

We now discuss the current limitations of our proposed evolutionary system and how future work may address these issues. In section 4, we presented a closed reaction network which was evolved to carry out a pre-specified task. We examined a second set of identical evolutionary simulations to examine whether alternative evolutionary trajectories could occur. In all of these experiments, we observed the convergence (from a phenotypic point of view) of the system towards reaction networks similar to Fig. 5.B. Moreover the evolutionary process would only involve a limited number of incremental steps in fitness. Based on these observations we may express that our proposed AC still lacks some evolvability and failed to display a "remarkable" evolution of complexity.

Some may argue that the current system is simply not complex enough and should incorporate more realistic features such as mass conservation, multidimensional space, a genetic code or advanced chemical kinetics. These complementary properties would certainly broaden the complexity of an already difficult and incomplete investigation. However there would be no guarantee of improved results, i.e., exhibiting a more interesting evolutionary growth of complexity: A first reason for this assertion is that the impact of environmental constraints on the evolution of complexity still remains to date an open question (Gershenson and Lenaerts, 2008). Secondly there is currently no theoretical framework that would lead to desired evolutionary growth of complexity with confidence.

We hypothesise that as long there is no such theoretical frameworks for the study of ACs, the latter would remain a black art approach where unexpected results are observed. Therefore a minimalist/simplified approach to ACs, where the system is still analytically tractable and examined using current mathematical methods, should be adopted. In keeping with this, our proposed AC contributed, to some extent, to the understanding of minimalist ACs and their evolutionary dynamics.

Future work would benefit from further theoretical research, it could involve the development of Formal Methods for Artificial Life (AL), this would facilitate the comparison of differing ACs. Such a formal approach to AL had already been initiated but was discontinuated (Grogono et al., 2003), an alternative could be to extend the AC formalism. Moreover the development of further analytical tools such as the Organization Theory, which was initiated by (Fontana and Buss, 1994a) and later enriched by (Dittrich and Speroni, 2007), would as well assist and contribute to the understanding and ultimately engineering of Complex Adaptive Systems using ACs.

Finally, although our approach may present some limitations in terms of evolutionary growth of complexity, it still demonstrated the successful evolution of closed reaction networks to accomplish a prespecified task. This may be regarded as an initial promising result demonstrating closure in CSNs and may suggest a novel method to engineer autonomous Complex Adaptive Systems for real-world applications.

6 CONCLUSIONS

We discussed how CSNs could be considered as subsets of closed reaction networks and the potential role of closure on CSNs evolution. We presented our evolutionary system: a multi-level concurrent Artificial Chemistry based on the Molecular Classifier Systems and the Holland broadcast language. The development of this multi-level and concurrent model was motivated to improve evolutionary stability, which was a key missing feature of the single-level MCS.bl. We then presented these different properties which permitted the systems robustness to be ameliorated. We finally presented an experiment in which a simple closed reaction network was successfully evolved and optimized to carry out a pre-specified task (signal amplification) whilst maintaining closure. This optimized signal-processing ability directly resulted from the evolved closure properties. We finally discussed the contributions and future directions of this work.

ACKNOWLEDGEMENTS

We acknowledge the ESIGNET (Evolving Cell Signalling Networks *in Silico*) Project funding (contract no. 12789).

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 Advances in Biologically Inspired Information Systems, volume 69, pages 165– 184. Springer.
- Decraene, J., Mitchell, G. G., and McMullin, B. (2008). Unexpected Evolutionary Dynamics in a String-Based Artificial Chemistry. *Proceedings of Artificial Life XI, MIT Press.*
- Dittrich, P. and Speroni, P. (2007). Chemical Organisation Theory. *Bulletin of Mathematical Biology*, 69(4):1199–1231.

- Fontana, W. and Buss, L. (1994a). The arrival of the fittest: Toward a theory of biological organization. *Bulletin of Mathematical Biology*, 56(1):1–64.
- Fontana, W. and Buss, L. (1994b). What Would be Conserved if "the Tape were Played Twice"? *Proc. of the National Academy of Sciences*, 91(2):757–761.
- Gershenson, C. and Lenaerts, T. (2008). Evolution of Complexity. *Artificial Life*, 14(3):241–243.
- Grogono, P., Chen, G., Song, J., Yang, T., and Zhao, L. (2003). Laws and life. In Proceedings of the 7th IASTED Conference on Artificial Intelligence and Soft Computing (ASC 2003), pages 158–163.
- Hogeweg, P. and Takeuchi, N. (2003). Multilevel selection in models of prebiotic evolution: compartments and spatial self-organization. *Origins of Life and Evolution of the Biosphere*, 33(4-5):375–403.
- Holland, J. (1976). Studies of the spontaneous emergence of self-replicating systems using cellular automata and formal grammars. *Automata, Languages, Development*, pages 385–404.
- Holland, J. (1992). Adaptation in natural and artificial systems. MIT Press Cambridge, MA, USA.
- J.Decraene (2006). The Holland Broadcast Language. Technical Report ALL-06-01, RINCE, Dublin City University. http://elm.eeng.dcu.ie/~alife/jd/ALL-06-01/.
- Kauffman, S. (1997). At Home in the Universe. Mathematical Social Sciences, 33(1):94–95.
- Machné, R., Finney, A., Müller, S., Lu, J., Widder, S., and Flamm, C. (2006). The sbml ode solver library: a native api for symbolic and fast numerical analysis of reaction networks. *Bioinformatics*, 22(11):1406–1407.
- McMullin, B., Kelly, C., and O'Brien, D. Multi-level selectional stalemate in a simple artificial chemistry. In *Proceedings of ECAL 2007*, LNCS. Springer.
- Ray, T. (1991). An approach to the synthesis of life. Artificial Life II, 10:371–408.
- Ray, T. (1992). Evolution, ecology and optimization of digital organisms. *Santa Fe.*
- Stadler, P. F., Fontana, W., and Miller, J. H. (1993). Random catalytic reaction networks. *Phys. D*, 63(3-4):378– 392.
- van Duijn, M., Keijzer, F., and Franken, D. (2006). Principles of Minimal Cognition: Casting Cognition as Sensorimotor Coordination. *Adaptive Behavior*, 14(2):157.