Enrichment of Interaction Rules in a String-Based Artificial Chemistry

Ciarán Kelly† and Barry McMullin and Darragh O'Brien
RINCE, Dublin City University, Glasnevin, Dublin 9, Ireland
†ciaran.kelly9@mail.dcu.ie

Abstract

In this paper, we discuss our reasoning and progress in adding a mapping between information and enzymatic function to our Molecular Classifier System (MCS). MCS takes a bottom-up approach to building artificial bio-chemical networks. Unlike Holland's LCS system, which it is loosely based on, MCS has no overt demarcation between *rules* and *messages*. In our previous work, we explored a version of this Artifical Chemistry which had an impoverished interaction scheme. While this system did present some interesting results, it had very limited potential for evolving greater complexity. We present here a mechanism for enriching the reaction rules used in our Artificial Chemistry. This mechanism is analagous to the folding of RNA to an enzymatically active form. To date, we have examined in detail the evolutionary trajectories of single reactors populated with this modified Artifical Chemistry and the results of this work are presented here.

Introduction

The field of Artificial Life was borne from the desire to understand "life" as a process, a complex dynamic system. Life, as we know it, evolved over billions of years to its current state. It is conjectured that there existed a "Last Universal Common Ancestor" (LUCA) from which every living thing descended (Forterre and Philippe, 1999). Granted, this ancestor lived a very long time ago, but to explain all the similarities between living things (e.g., DNA as a genetic molecule, the almost unique genetic code between DNA and protein etc.), the case for the existence of LUCA is strong. But, while LUCA would be an ancestor of all currently living things, it was surely not the *first* living thing. A key aspect of origin-of-life research therefore focuses on the time before LUCA—who or what were LUCA's ancestors?

The initial explorers of Artificial Life, such as Von Neumann (Von Neumann and Burks, 1966), examined the qualitative properties of life: what enables us to say that this thing is alive and this thing is not? In certain senses, the chemistry of life is now well understood. Genetic theorists since Mendel have understood the basic hereditary mechanics of life. Biochemists can already build arbitrary strands of DNA containing whatever nucleotide sequence they wish. The human genome project has sequenced and catalogued every gene in the human genome. Why, then, can we not just put together some carefully chosen pieces of DNA to create new life forms? Indeed, it would only need to be done once—the newly synthesized creatures could presumably reproduce, self-repair and evolve to cope with environmental perturbations. The answer, of course, is that there is much more to life than the sequence of monomers on a polymer, or even of genes on a chromosome. Artificial Life is, in part, the study of what that "more" is: how it can be characterised, where it comes from, and how it could be exploited it to advance human technology. We suggest that a return to the first principles of origin-of-life research may help us understand these fundamental qualities in a more abstract way.

In this context, a return to first principles means not so much replaying the tape of life (Gould, 1989), but rather examining, in detail, key stages in the evolution of living entities from non-living matter in order to abstract some *rules* that describe what life really is. Our approach is to build an Artificial Chemistry (Dittrich et al., 2001) that abstracts away from some of the "*chemical*-specific" problems, and focuses more on "*organisation*-specific" problems. That is not to say that the chemical problems are trivial, but that we want to separate the study of the organisation of simple life-forms from the specific requirments of terrestrial carbon-chemistry—"life-as-it-could-be" rather than "life-as-we-know-it" (Langton, 1989). Bedau has previously discussed the nature of life and presented the idea that life is an emergent macro-level property of systems rather than being dependant on the composition of the micro-level entities that make up these systems (Bedau, 1996, 1999). This is an important idea because it suggests a case for the exploration of digital life—life *in silico*. *In silico* experiments allow the direct study of the emergent properties of life, without the need to first solve the chemical problems that "life-as-we-know-it" has already solved.

Of course, before we can approach anything like *in silico* life, we need to be somewhat careful about how we define life in the first instance. We adopt here the definition proposed by Maynard Smith and Szathmáry (Maynard Smith

and Szathmáry, 1997)—an entity is alive if it has the properties of multiplication, variation and heredity, or if it is descended from entities which exhibit those properties. Populations of such entities which are forced to compete with one another will undergo Darwinian natural selection. This definition could clearly be applied to digital life, as it enforces no requirements on the material substance of life.

Bedau et al. (Bedau et al., 2001) have presented a list of open problems in Artificial Life. Our research is focussed on the exploration of the "Transition to Life, *in silico*", which was one of the open questions identified. Protocells are hypothesized as a transitional phase in the evolution of the biosphere (Maynard Smith and Szathmáry, 1997). In previous work we have constructed an Artificial Chemistry as a platform upon which to investigate the evolution of "computational" functionality in protocells (McMullin et al., 2007a,b). We started with a minimal "template-replicator world" in which there was only one level of Darwinian actor (the replicating "molecule"). This model incorporated the notion of unlimited heredity achieved through (catalysed) template replication of indefinite length polymers. In the simplest case we considered molecules which could act as "self-replicases"—a form of degenerate, one-element, hypercycle (Eigen and Schuster, 1977).

That work examined the inclusion of an elementary form of mutation. Molecular replication was made imperfect, with a fixed error rate per monomer (thus the molecular-level replication error rate increases with the length of the molecule).

We further introduced a simple rule for enzymatic coupling between different species (so that one species can function as replicase for another species as well as itself). This was deliberately made asymmetric. This introduced the possibility of exploitation between species. Even under the condition of hyperbolic growth[1], this allows effective displacement of a "host" species by a new "facultative parasitic" species; and, under the conditions of the model, this can happen repeatedly. In this particular model, this leads to the somewhat counter-intuitive effect of systematic, macro-evolutionary, *deterioration* in "intrinsic fitness" (as measured by replication *fidelity*).

To investigate the more interesting phenomenon of multi-level selection, populations of these molecules were injected into externally provided protocells, where protocell reproduction (by binary fission) is driven by molecular replication. By fixing the size of the protocell population, we imposed a distinct process of Darwinian selection at the higher hierarchical level of the protocell. The protocell level of selection is governed by the molecular level selection dynamic which still occurs within each protocell. We showed that the protocell level selection did effectively control parasitic exploitation at the molecular level; however, the molecular level selection is still effective in preventing positive evolution in the opposite direction (toward higher molecular-level replication fidelity). The result was a rather robust evolutionary "stalemate" in which the selectional dynamics at the two interacting levels were, in effect, precisely counteracting each other.

The system presented in these works was, of course, a radically simplified version of the phenomena that occur in real chemistry and biology. Its purpose was not to directly model such real systems. Rather it was presented as a deliberately minimal system which already illustrated how complex and counter-intuitive the evolutionary behaviour of such systems could be; but also, how the evolution could, indeed, be dramatically altered by the interaction between multiple levels of selection.

The broader intention of the current work is to develop a minimal abstract framework for understanding the evolutionary emergence of "computation" or, at least, coordinated signal processing and control, in protocellular systems. Presumably, any interesting molecular level computation must rely on a diversity of chemical species; but all of these in turn must be "replicated", directly or indirectly, to support protocell level reproduction.

The work presented in this paper addresses our progress towards the incremental widening of the repertoire of molecular interactions. Again, the broader context of this work is to explore the impact that these new interaction schemes will have on the multi-level (protocell-based) selection model, though we do not discuss such hierarchical selection in this paper.

The Molecular Classifier System

We propose a highly simplified Artificial Chemistry loosely based on John Holland's Learning Classifier Systems (Holland, 2006; Holland and Reitman, 1977), which we call the Molecular Classifier System (MCS).

The operation of our system depends on a population of "molecules", which take the form of binary strings. Each molecule has an informational structure (primary structure, or monomer sequence) and an enzymatic function ("folded" or secondary structure, or "shape"), as inspired by the ribozymes of the RNA world hypothesis (Joyce, 1991). The model also contains a rule-set which determines the enzymatic action to take, given a particular molecule. Our artificial protocells are then crudely modelled as containers for a dynamic mix of these molecules, which continuously interact and exert enzymatic actions on each other. This "informational chemistry" might then be evolved to realise some particular computation—provided that it is simultaneously capable of sustaining its own dynamic organisation. In particular, this informational or computational sub-system must grow (in absolute number of molecules) and divide in co-

[1] There is, of course, a large body of prior literature on replicator selection dynamics. We omit any extensive review here, in the interests of brevity; but (Szathmáry and Maynard Smith, 1997), for example, includes a comprehensive bibliography.

ordination with overall cell reproduction.

For the purposes of the specific model to be discussed here, the only supported enzymatic function is, by design, to make an error-prone bit-wise copy of the primary, informational, structure of the bound, substrate, molecule; that is, a *replicase* function. More specifically, if a particular molecule has the ability to bind to molecules with the same molecular structure as itself, it will effectively be able to function as a *self*-replicase.[2]

This restriction of enzymatic function to replication only is, of course, a radical simplification of any real chemistry; and, further, is a significant limitation of the potential dynamics. Additional significant simplifications are that all reaction *rates* are equal, and replication error rate (per monomer) is constant. Nonetheless, we suggest that it should be useful to fully understand the variety of selectional dynamics that are possible even in this simplified case first, before introducing the additional complications of more complex and varied enzymatic function, reaction rates etc. It is, of course, a longer term goal of the research to systematically re-introduce these more complex and realistic properties.

Model Setup

Our basic template-replicator world consists of a finite number of strings (polymers) drawn from a binary alphabet. The dynamics consists of a simple loop in which one random string is chosen as a replicase and a second as a template. If the replicase is determined to "match" the template (via a molecular transformation to be discussed later), then it "binds" to it, and replicates it, with a specific bit-wise error-rate. Another molecule is chosen at random and is replaced by the new molecule. It should be noted that there is no specific modelling of dilution flux (ϕ). If the replicase does not bind then the interaction is considered to be elastic.[3]

For the purposes of analysis we consider any specific pair of molecular species to give rise to a "binary replicase-reaction network", i.e., a network comprising the two distinct replicase-reactions that can occur between these species depending on which one functions as the enzyme and which as the substrate (or replication-template, as replication is, for the moment, the only supported enzymatic function).

In our previous work (McMullin et al., 2007a), we presented an approximate analysis of some particular binary replicase-reaction network using an appropriate set of ordinary differential equations (ODE). This allowed predictions of the concentration dynamics of a flow-reactor populated by a single pair of molecular species. We now extend that analysis to systematically examine and classify *all* possible binary replicase-reaction networks in this general type of model chemistry.

Binary Replicase-Reaction Networks

In order to construct the ODE representations of the reaction kinetics, we first derive a set of Binary Replicase-Reaction Networks classes. These classes are generic in the sense that any MCS-like system can be represented by them. They are constructed by considering the reaction kinetics: two molecules are chosen and a reaction is attempted. If we represent this scheme as a logical truth-table, we can easily enumerate and classify all possible such networks. The truth-table is constructed by considering two distinct molecular species. Each molecule may, or may not be a self-replicase—i.e., it may or may not be able to bind to copies of itself. At the same time, each molecule may or may not be able to act as a replicase for the other molecule—i.e., it may or may not be able to bind to copies of the other molecule. Taking all of these possibilities into consideration, there are 16 possible truth-tables which represent every possible combination of two molecules and two reaction rules. Allowing for certain symmetries and equivalences, these reduce to set of 10 properly *distinct* tables.

Any specific binary replicase-reaction network can be represented as follows:

$$\begin{bmatrix} (XX) & (XY) \\ (YX) & (YY) \end{bmatrix}$$

where a 1 in the (XX) position means "X is a self-replicase" and a 0 means that "X is not a self-replicase". Similarly, a 1 in the (XY) position means "X can replicate Y" and a 0 means that "X can not replicate Y".

In (McMullin et al., 2007a) we showed that it was possible to formulate an approximate differential equation model of this system. We consider two species (X and Y). Taking their respective relative concentrations as x and y, these are also the probabilities of choosing an instance of either species at random. As an example, assume X is a self-replicase. The probability of choosing two X molecules and the offspring displacing a Y molecule is evidently x^2y. Thus, the growth rate[4] of x is given by:

$$\dot{x} = x^2 y$$

Of course, this a deterministic approximation using continuous concentration values; real implementations will have discrete numbers of each molecular species and the dynamics will be stochastic. Nontheless, this ODE analysis should provide a qualitative baseline for the expected dynamic behaviour, at least as long as significant numbers of each species are present.

[2] To our knowledge, no real RNA self-replicase has yet been identified, though the conjectured existence of such molecules plays a core role in the RNA-world hypothesis.

[3] The system is therefore generically a "catalytic reaction network" in the sense of (Stadler et al., 1993).

[4] In this and subsequent equations there is an implicit multiplicative constant, effectively setting the time scale. This has been arbitrarily taken as unity.

By applying this method and discarding all the reactions which have zero effect on the concentrations, we can convert the truth-tables into differential equations. For this initial analysis we are neglecting mutation, so y, the concentration of Y, will trivially be $(1-x)$, and $\dot{y}$ will be $(1-\dot{x})$. In each case we therefore explicitly provide just the expression for $\dot{x}$.

The following terminology will be used when presenting the binary replicase-reaction networks:

- ***Sterile*** molecules can neither replicate themselves nor be replicated by another molecule.
- ***Self-Replicase*** molecules can replicate themselves, but cannot be replicated by another molecule.
- ***Obligate Parasite*** molecules cannot replicate themselves, but *can* be replicated by another molecule.
- ***Facultative Parasite*** molecules can both replicate themselves and be replicated by another molecule.

Once the relevant differential equations have been extracted for each binary replicase-reaction network we can make predictions about the flow-reactor dynamics that each network gives rise to.

Class 0:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

The two molecular species are ***Sterile***. The ODE representing growth rate for both species is therefore, trivially, 0.

Class 1:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = 0$$

The two molecular species are ***Facultative Parasite***s. In this way, there is full cross-catalysis between the molecules, but since neither molecule has a distinct advantage over the other, the growth rate ODE for both species is again 0.

Class 2:

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = -y^2x \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = x^2y$$

One molecule is a ***Self-Replicase*** and the other is a ***Sterile*** molecule. As would be expected, the ODE analysis states that the ***Self-Replicase*** will displace the ***Sterile*** molecule.

Class 3:

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = y^2x \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = -x^2y$$

One molecule is an ***Obligate Parasite*** and the other is a ***Sterile*** molecule. The ODE here shows that the ***Obligate Parasite*** will displace the ***Sterile*** molecule. As the concentration of the ***Sterile*** molecule decreases, so too does overall reaction rate (in the limit, when the final ***Sterile*** molecule is eventually eliminated, there will be no further reactions at all).

Class 4:

$$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = 0 \quad \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 0$$

One molecule is a ***Self-Replicase*** and the other is an ***Obligate Parasite***. In this case, a 0 growth rate is indicated by the ODE. This is explained by the fact that, by our reaction kinetics, the ***Self-Replicase*** will replicate a copy of itself exactly as often as replicating a parasite molecule.

Class 5:

$$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} = -x^2y - y^2x \quad \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = x^2y + y^2x$$

One molecule is a ***Facultative Parasite*** and the other is a ***Sterile*** molecule. Again, it is easy to see that the ***Sterile*** molecule will be completely displaced, but the reaction rate then continues at the maximum level (albeit with no further change in concentration).

Class 6:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = x^2y \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = -y^2x$$

One molecule is a ***Self-Replicase*** and the other is a ***Facultative Parasite***. The ODE analysis for this situation shows that the ***Facultative Parasite*** will completely displace the ***Self-Replicase***. This was the generic case considered in detail in (McMullin et al., 2007a).

Class 7:

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = -x^2y \quad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = y^2x$$

One molecule is a ***Facultative Parasite*** and the other is an ***Obligate Parasite***. In this case the ***Facultative Parasite*** will always displace the ***Obligate Parasite***.

Class 8:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -x^2y + y^2x$$

Both molecular species are ***Obligate Parasite***s. This essentially means Class 8 networks are two-component hypercycles. Neither species can replicate itself but each can catalyse the replication of the other. The concentration of each species will therefore be maintained at exactly equal levels.

Class 9:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = x^2y - y^2x$$

Both molecular species are independent ***Self-Replicase***s; the "survival of the common" applies, so that whichever initially achieves a significantly higher concentration will then completely displace the other. Again, this case was detailed in (McMullin et al., 2007a).

Molecular Binding Rules

Bit-Wise Substring Binding

The binding rules that the system uses are now discussed. In (McMullin et al., 2007a), we explored perhaps the simplest

binding-rule—bit-for-bit sub-string matching with no distinction or mapping between informational (primary) structure and enzymatic (secondary) structure. If the replicase exactly matched the template in sequence, it was assumed to bind to it. This meant that the selection of binary replicase-reaction networks that could be observed was considerably smaller than the total number of networks. The first thing to notice is that every molecule is a self-replicase, and for two molecules of the same length, the only possible replicase-reaction network is the one named "Class 9" above, since it is logically impossible for two non-similar strings of the same length to be sub-strings of each other. The interesting results we achieved during that work were due to what we now call "Class 6" networks. As described above, a "Class 6" network is a network consisting of a ***Self-Replicase*** molecule and a ***Facultative Parasite***. Since the binding rule used was "bit-wise substring", the only possible way for a "Class 6" network to emerge from a mixture is if a lengthening mutation occurs during the copy of one of the templates such that the new molecule contains the "parent" molecule's structure as a sub-string.

We showed that in all cases (modulo statistical fluctuations), the "parasite" could invade a population of altruistic hosts—a pathology characterised by progressive lengthening of the average molecular string length over macro-evolutionary time. Of course, the key difficulty here is that since we are using a per-bit mutation rate, the "per-molecule" mutation rate will increase with molecular length. In a single-reactor the effect manifested as a reduced reaction rate—a lengthening of the time between successful reactions. This was due to the difficulty in finding two matching molecules to react: as mutation rate increases, so does the relative population of mutants. However, our protocell level experiments yielded an even more distinctive result. Once we added the new level of selection, at the protocell level, the parasitic behaviour of the interacting molecules was effectively halted. This was a direct consequence of hierarchical selection. The explanation is that lineages of protocells which have higher reaction rates will do better, on average, than those with lower reaction rates, since the faster a protocell can grow, the more often lineages of such protocells will undergo cellular division.

The actual effect of the "multi-level" selection was that of "selectional stalemate"—average molecular length neither increased nor decreased in the protocell model. The internal molecular dynamics ensured that under no circumstances could a "shorter" (relatively speaking) molecule come to dominate any protocell—"shorter" molecules are parasitised by "longer" ones, *"and so on, ad-infinitum"*. This tendency towards longer molecules leads to an increased mutational load on the individual protocell, and a corresponding decrease in reaction rate for that protocell.

The net effect is that, over a wide range, this system can be initialised with a protocell population with any arbitrary dominant molecular length; and the population will then remain dominated indefinitely by protocells which are individually dominated by this initial specific molecular species. Evolution toward protocells dominated by longer molecules will be prevented by the protocell level selection; and evolution toward protocells dominated by shorter molecules will be prevented by molecular level selection.

More Flexible Binding

The previous results summarised above are determined by the fact that there were really only two possible binary replicase-reactions possible with that implementation.

By opening up more reaction network possibilities, it was predicted that one could increase the variety of the system behaviour. One biochemically inspired method to go about this was to implement a mapping mechanism, similar to the folding of RNA to an enzymatically active form, that re-opens the possibility to have all possible binary replicase-reaction networks. We decided that the most incremental approach was to process the molecules in chunks of two bits, and to map these pairs into some secondary, functional, alphabet. This pair-wise processing allows for a secondary alphabet of 4 symbols. This alphabet, and the coding scheme is defined such that all 10 distinct binary replicase-reaction networks are realisable. Table 1 and Table 2 offer a comparison between the new and previous coding schemes.

Table 1: Previous Coding Scheme

chunk	function	description
0	L	match literal '0'
1	H	match literal '1'

Table 2: Enriched Coding Scheme

chunk	function	description
00	L	match literal '0'
01	L	"
10	H	match literal '1'
11	H	"

It is obvious that a molecule (bitString) which is processed by the scheme given in Table 2 will result in a functional string that is shorter than if the same molecule was processed by the scheme given in Table 1. In this version of MCS, the molecular binding rule is still bit-for-bit sub-string matching, but now, the matching happens between the functional string derived from Table 2 and the molecular bit string of the substrate molecule. In this system therefore, all 10 distinct binary replicase-reaction networks can be instantiated.

Conjecture (and Refutation) Let us assume a single-reactor with this new MCS chemistry inside. If we seed this reactor with a ***Self-Replicase*** molecule, we would once again expect this ***Self-Replicase*** to remain dominant in spite of mutation, until at some point a ***Facultative Parasite*** arises in the population and displaces it. An examination of the ODE for each of the binary replicase-reaction networks when focussing on a parasite attempting to invade the population would provide an understanding of the reasoning behind this form of hypothesis. For each reaction network we considered the growth potential of a molecule which, in combination with the seed species—the species with the highest concentration—forms such a binary replicase-reaction network. However, the ODE model would suggest that the only network where the new molecule could have a reliable advantage is that observed in a "Class 6" network. This is the same network class that allowed the parasitic take-over in the scheme represented by Table 1.

Predicted Results Our hypothesis was based upon an assumption that the reaction classes could be understood individually and that the evolutionary trajectory of a particular reactor could be predicted based on the reaction dynamics of the class that dominated the reactor. Furthermore, the ODE analysis of the reaction networks led us to believe that, once a reactor was dominated by a self-replicator, only one type of displacement event could reliably take place, namely "Class 6" ***Facultative Parasite*** driven displacement. We noted that due to the nature of the primary / secondary alphabet mapping, it was now possible to get a parasite that was shorter than its host. Further work is necessary to fully evaluate how things would be different in a multi-level, hierarchical selection situation. From our previous work (McMullin et al., 2007a), we know that hierarchical selection applies fitness pressure in the direction of better reaction rates, which results in shorter molecules. One could predict that with a parasite that is shorter than its host, the molecular level of selection and the cellular level of selection might become aligned given the correct initial conditions.

Observed Results We predicted above that a reactor, if seeded with a large number of a given self-replicase, would remain dominated by that species, at least for some reasonably extended period of time (i.e., until a facultative parasite results from mutation). In fact, it turned out that even with a low mutation rate (0.01 mutations per bit copied), a reactor seeded with a dominant replicator with some mutational copies will always result in the rapid displacement of the original seed species by a diverse variety of other species, none of them present in individually large concentrations. This result is shown in Figure 1, summarising 10 independent runs of the model.

If we further analyse one of these individual runs in more detail, we can observe that the second part of the hypothesis, that is, that the decay in concentration of the seed species would be due to the arrival of a ***Facultative Parasite***, is also flawed. Figure 2 shows the results obtained by graphing the concentration of molecules for each binary replicase-reaction network, relative to the initial seed, self-replicase, molecule. This shows that, rather than a Class 6 network emerging, Class 1 and Class 4 networks are the most prevalent contributing factors to the dilution in concentration of the seed species.

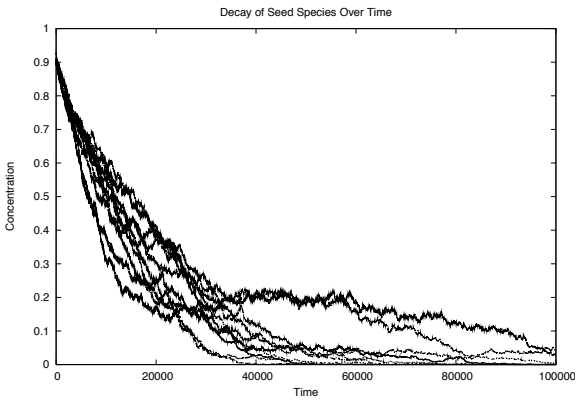

Figure 1: Concentration Decay of Seed Species

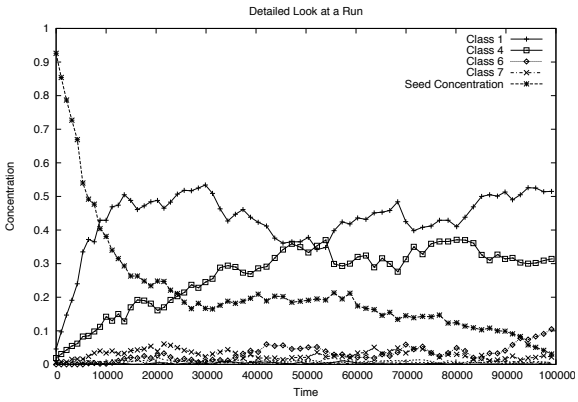

Figure 2: Detailed Analysis of a single Run

Analysis As a first step in understanding what was going on, we reviewed the earlier, ODE based, classification into 10 distinct binary replicase-reaction network classes. That analysis had suggested that the only way a molecule could reliably displace a currently dominant host was if that molecule was a ***Facultative Parasite*** of the host—they share a "Class 6" relationship. The ODE analysis for all other equivalence classes suggested that there could be no "invasion-from-rarity" displacement event. We tested these predictions in isolation, by seeding reactors with only two species of molecule and with mutation disabled. In all cases, the behaviour was as predicted by the ODE model. The only

set of experiments which led to selective displacement of the seed species were those involving displacement by "Class 6", ***Facultative Parasite***s. Of course, this does not explain the dynamics observed in Figure 2.

Further analysis has shown that the fate of these reactors is far more complicated. We have observed that "Class 1" mutants lead to the slow dilution of the seed species concentration. This dilution is a consequence of the asymmetric way in which mutation is applied. Our ODE analysis implicitly assumed that the rate of mutant generation from replication of the seed species would be balanced by an equal but opposite back-flow of mutant copies, since mutation rates were constant. Some brief analysis was carried out which highlighted the fact that for a given mutant offspring of the seed species, there are many possible mutational copies that could arise, compared with the one single mutational pathway back to the *specific* master species that gave rise to it. This asymmetric mutation pattern meant that there could be a consistent, nett mutational flow from the seed species into the collection of nearby mutants[5]. Further analysis showed that "Class 1" mutants arose most frequently in the nearby mutational neighbourhood of the seed species, and thus would be expected to arise most often.

Our experiments have clearly shown that the interplay between ***Facultative Parasite***s, ***Obligate Parasite***s and "Class 1" 'promiscuous' mutants can cause the decay of the seed species concentration. Upon further analysis, "Class 1" mutants arise initially, gain ground against the seed species and begin the one-way dilution of its concentration. Once the seed species begins to lose dominance the mutants become evermore involved in successful reactions so that their combined effect cannot be discounted any longer.

Conclusion & Future Work

In our previous published work, we demonstrated a system, MCS, consisting of two interdependent, opposing levels of selection. The macro-evolutionary outcome was that of selectional stalemate—the selectional pressures at each level exactly balanced. These previous experiments showed that for any appropriate set of initial conditions, the system would stabilise exactly where it began and evolutionary growth would essentially cease. In this paper however, we present some modifications to the chemical reaction rules of the MCS which allow for a richer set of interactions.

Initially we described our efforts to enrich the rules governing "chemical" interactions in the MCS. We accomplished this by taking inspiration from biochemistry, i.e., folding of RNA into an enzymatically active form. We added the concept of a more flexible binding by adding the simplest secondary, functional, structure to each molecule. This opened up the possibility of having molecular interactions that did not rely on exact substring matching between molecular bit-strings.

Our hypothesis was that a reactor could be initialised with a seed species of dominant concentration, and that *that* seed species would remain at dominant concentration until it was displaced by a "Class 6" ***Facultative Parasite***. However, our experiments have proved to be only partially successful in supporting this hypothesis. We found that the concentration of the seed species would decay, but that this decay was not necessarily associated with the arrival of a ***Facultative Parasite***. We believe that further experimentation with single-reactor dynamics is required before we attempt any experiments with a system which implements hierarchical selection.

Our work addresses some of the issues surrounding the understanding of "life-as-it-could-be" rather than what is currently examinable *in-vitro*—"life-as-it-is". Our work takes a bottom-up (ie. from level zero) approach to the simulation of evolutionary systems which appear to display obvious dynamics which may have been taken for granted until now.

[5]Mutants are considered to be "nearby" if they occur within a reasonably low *levenshtein* (string-edit) distance from the seed species

Acknowledgements

This work has received financial support under EU FP6 Integrated Project PACE (contract number 002035). The authors would also like to thank George G. Mitchell, James Decraene and Pietro Speroni di Fenizio for helpful discussions at the aLife lab that shaped the core material of this work.

References

Bedau, M. A. (1996). The nature of life. *The Philosophy of Artificial Life*, pages 332–357.

Bedau, M. A. (1999). Can unrealistic computer models illuminate theoretical biology. *Proceedings of the 1999 Genetic and Evolutionary Computation Conference Workshop Program*, pages 20–23.

Bedau, M. A., McCaskill, J. S., Packard, N. H., Rasmussen, S., Adami, C., Green, D. G., Ikegami, T., Kaneko, K., and Ray, T. S. (2001). Open Problems in Artificial Life. *Artificial Life*, 6(4):363–376.

Dittrich, P., Ziegler, J., and Banzhaf, W. (2001). Artificial Chemistries - A Review. *Artificial Life*, 7(3):225–275.

Eigen, M. and Schuster, P. (1977). The Hypercycle, A Principle Of Natural Self-Organization. *Die Naturwissenschaften*, 64(11):541–565.

Forterre, P. and Philippe, H. (1999). The Last Universal Common Ancestor (LUCA), Simple or Complex? *Biological Bulletin*, 196(3):373–377.

Gould, S. J. (1989). *Wonderful Life*. Penguin Group, London.

Holland, J. H. (2006). Studying Complex Adaptive Systems. *Journal of Systems Science and Complexity*, 19(1):1–8.

Holland, J. H. and Reitman, J. (1977). Cognitive systems based on adaptive algorithms. *ACM SIGART Bulletin*, pages 49–49.

Joyce, G. F. (1991). The rise and fall of the RNA world. *The New Biologist*, 3(4):399–407.

Langton, C. G. (1989). Artificial Life. In Langton, C. G., editor, *Artifical Life*, pages 1–47. Addison-Wesley Publishing Company, Inc., Redwood City, California. Proceedings of an interdisciplinary workshop,Los Alamos, New Mexico, September, 1987.

Maynard Smith, J. and Szathmáry, E. (1997). *The Major Transitions in Evolution*. Oxford Press.

McMullin, B., Kelly, C., and O'Brien, D. (2007a). Multi-Level Selectional Stalemate in a Simple Artificial Chemistry. *ECAL 9, Lisbon, Portugal*.

McMullin, B., Kelly, C., and O'Brien, D. (2007b). Preliminary Steps toward Artificial Protocell Computation. *International Conference on Morphological Computation, Venice, Italy*.

Stadler, P., Fontana, W., and Miller, J. (1993). Random catalytic reaction networks. *Physica D*, 63(3-4):378–392.

Szathmáry, E. and Maynard Smith, J. (1997). From Replicators to Reproducers: the First Major Transitions Leading to Life. *Journal of Theoretical Biology*, 187:555–571.

Von Neumann, J. and Burks, A. W. (1966). *Theory of Self-Reproducing Automata*. University of Illinois Press Champaign, IL, USA.