Multi-Level Selectional Stalemate in a Simple Artificial Chemistry

Barry McMullin, Ciarán Kelly, and Darragh O'Brien

aLife Lab, RINCE, Dublin City University
`barry.mcmullin@dcu.ie`
WWW: `http://www.eeng.dcu.ie/~alife/`

Abstract. We describe a simple artificial chemistry which abstracts a small number of key features from the origin of life "replicator world" hypotheses. We report how this can already give rise to moderately complex and counter-intuitive evolutionary phenomena, including macro-evolutionary deterioration in replication fidelity (which corresponds to intrinsic replicator fitness in this model). We briefly describe the extension of this model to incorporate a higher, protocell, level of selection. We show that the interaction between the two levels of selection then serves to control parasitic exploitation at the molecular level, while still significantly constraining accessible evolutionary trajectories at the protocell level. We conclude with a brief discussion of the implications for further work.

1   Introduction

It has been argued that the most significant events in the evolution of the biosphere have been a relatively small number of so-called "major transitions" [6]. These are typified by the emergence of a new "level" of Darwinian actor, and thus a new process of Darwinian evolution. Thus, for example, there may have been an epoch early in the origin of life in which replicating molecules directly functioned as Darwinian actors (e.g., an RNA World [4, 5]); but these independently replicating molecules subsequently became subsumed into some form of cellular containers or "protocells". The evolutionary fate of any molecular level species then became contingent on that of the cellular aggregate of which it was a component. Similarly, there was another major transition between unicellular and multi-cellular organisms. In the latter, the evolutionary fate of the individual cells became subjugated to that of the multi-cellular aggregate.

A critical feature of any such major transition is the interplay between the two levels of selection. In particular, selection at the lower level is generally expected to favour "selfishness" on the part of the lower level actors [1].[1] However, as the evolutionary future of these entities becomes conditional on the success of the new, higher level, entity, this selfishness at the lower level may necessarily have

---

[1] Or, more precisely, on the part of *lineages* of these actors [7].

to be moderated in favour of more "co-operative" interactions which better serve the higher level, composite, entity.

We present here a set of artificial chemistry style [2] investigations of this phenomenon.

We start with a minimal "template-replicator world" in which there is only one level of Darwinian actor (the replicating "molecule"). The model incorporates the notion of unlimited heredity achieved through template replication of indefinite length polymers. This is taken to be dependent on catalysis by a suitable replicase. In the simplest case we consider molecules which can act as "self-replicases" (a form of degenerate, one-element, hypercycle [3]). We note and exhibit the distinctive hyperbolic selection dynamics of such systems ("survival of the common").[2]

Molecular replication is then made imperfect, with a fixed error rate per monomer (and thus a molecular-level replication error rate increasing with the length of the molecule). We characterise the effect of this on the distribution between the dominant and mutant molecular species.

We next introduce 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 is deliberately made asymmetric. This introduces the possibility of exploitation between species. Even under the condition of hyperbolic growth, this allows effective displacement of a "host" species by a new "parasitic" species; and, under the conditions of the model, this can happen repeatedly. In this particular model, this has the somewhat counter-intuitive effect of leading to systematic, macro-evolutionary, *deterioration* in "intrinsic fitness" (as measured by replication *fidelity*).

Finally, we extend the model to place the molecules in simple "protocells", where protocell reproduction (by binary fission) is driven by molecular replication. With a fixed protocell population limit we now have a higher level Darwinian dynamic at the protocell level which interacts with the molecular level selection dynamic which still occurs within each protocell. We show that the protocell level selection does effectively eliminate 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 is a rather robust evolutionary "stalemate" in which the selectional dynamics at the two interacting levels are, in effect, precisely counter-acting each other.

We conclude with a brief discussion of the implications of this study.

2 Basic Template-Replicator World

The basic template-replicator world consists of a finite number of strings (polymers) on a binary alphabet. The dynamics consists of a simple loop in which

---

[2] 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 [8], for example, includes a comprehensive bibliography.

one random string is chosen as a replicase and a second as a template. If the replicase exactly matches the template in sequence, then it "binds" to it, and replicates it. A random existing string is removed to make room for the new string. If the replicase does not bind (i.e., does not match in sequence) then the interaction has no effect.

Note that, by design at this point, there are no replication errors, and indeed no other intrinsic fitness differences between different molecular species.

Clearly, any species acts as a replicase for itself and no other; which is to say, there is a form of replication, or auto-catalysis. However, because of the requirement for two molecular instances to interact to achieve replication, the autocatalytic reaction scheme is of the form $2A \rightarrow 3A$ rather than the more usual $A \rightarrow 2A$. This has a significant effect on the dynamics. The intrinsic growth is not exponential, but hyperbolic. The result is effectively a form of positive frequency dependent selection, in which any species with above average representation will experience a direct positive feedback. In the absence of opposing intrinsic fitness differences, any species which initially achieves such an above-average representation will quickly become dominant; and will then be able to maintain that dominance even against invasion by species with much higher intrinsic fitness. This is sometimes referred to as "survival of the common" [8].

We can formulate an approximate differential equation model of this system by considering just 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. If we choose two X molecules they will successfully replicate; and if the offspring displaces a Y molecule, there will be a net increase in x. The probability of this event on any given interaction is evidently $x^2 y$. Conversely, if we choose two Y molecules and then displace an X, this will give a decrease in x, with probability $y^2 x$. All other interactions leave the concentrations unchanged. Thus, the rate of change[3] of x is given by:

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

Given that $x + y = 1$ (by definition), then we have:

$$\dot{x} = x^2(1-x) - (1-x)^2 x$$

A simple insight into the behaviour is gained by plotting the expression on the right against x. This is shown in Figure 1. As can be seen, if x ever becomes greater than 0.5 then its rate of change will be positive, and it will keep on increasing until it completely displaces y (i.e., reaches concentration 1). Of course, by symmetry, exactly the same is also true of y. While there is a theoretical steady state with $x = y = 0.5$ this is clearly unstable.

The behaviour in a real system will obviously depend on the initial condition, and statistical fluctuations in a finite population. But even if initialised with $x = y = 0.5$ we expect that as soon as there is any significant fluctuation in

---

[3] In this and subsequent equations there is an implicit multiplicative constant, effectively setting the time scale. This has been arbitrarily taken as unity.
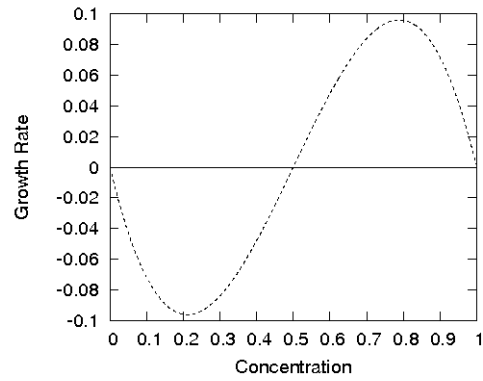
Fig. 1. Selection under hyperbolic growth.

favour of either species, that species will then quickly take over.[4] For illustration, data from one concrete experimental run illustrating this is shown in Figure 2.
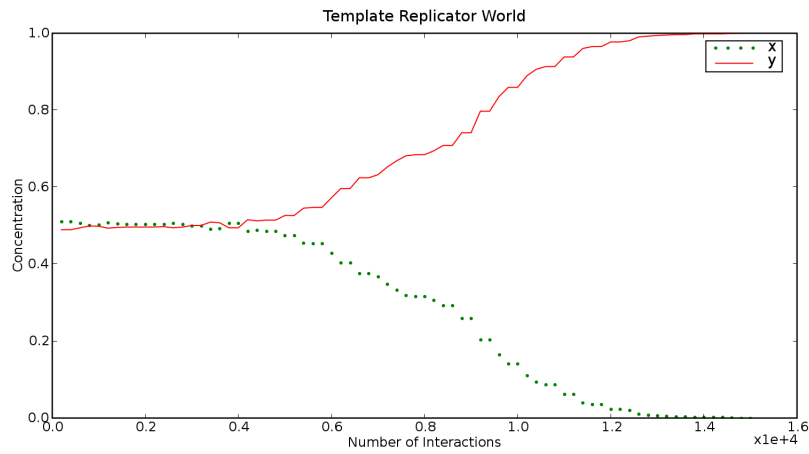

Fig. 2. Example of Template-Replicator World Experiment

---

[4] Note that this behaviour is in marked contrast to systems of simple exponential replicators, in which, in the absence of fitness differences, relative concentrations will undergo only slow random drift.

Replicase	Template	Product	Replaces	Net Effect	Probability
X	X	X	X	–	
X	X	X	Z	increase x	$x.x.(1-M).z$
X	X	Z	X	decrease x	$x.x.M.x$
X	X	Z	Z	–	
X	Z	–	–	–	
Z	Z	–	–	–	
Z	X	–	–	–	

Table 1. Interactions under Molecular Mutation

3 Effect of (Molecular) Mutation

In the simple template-replicator world, the population will quickly be taken over by whichever species initially happens to achieve a greater than average representation. In the absence of replication error, this species will simply fixate at concentration 1.0. However, we now wish to consider the effect of replication error or molecular "mutation".

Qualitatively, the frequency dependent advantage of a dominant sequence is expected to be strong enough to maintain that dominance even when subject to comparatively high replication error rates. We can crudely analyse this situation as follows. Let X (concentration x) denote the dominant species; and let Z (concentration z) denote the collective population of all other (non-X) sequences. We will assume that we can neglect the replication of Z molecules.[5]

Let the per-bit error rate be m. If X is n bits long, then the per-molecule error rate will be:

$$M = 1 - (1-m)^n$$

We can now summarise the outcomes of the possible interactions as shown in Table 1, yielding the approximate differential equation:

$$\dot{x} = (x^2(1-M)z) - (x^3 M) = x^2((1-M)z - Mx)$$

Given that $x + z = 1$ this yields:

$$\dot{x} = x^2(1 - x - M)$$

We can find the stable states by setting $\dot{x}$ to zero. Neglecting the case of $x = 0$ this leaves one fixed point with:

$$1 - x - M = 0$$

that is:

$$x = 1 - M$$

---

[5] This is based on assuming that X is of sufficient length that the mutated sequences are diverse and almost never find an identical sequence to bind to. This could be analysed more precisely, but this coarse simplification is adequate for our purposes here.

Replicase	Template	Replace	Effect
x	x	y	increase x
x	y	x	decrease x
y	y	x	decrease x

Table 2. Parasitic Invasion

Thus, the steady state population structure is with a proportion $(1-M)$ of the dominant species and M of diverse mutants. Note that, because of the (assumed) diversity in the mutant population, and the frequency dependent advantage of dominance, a particular species may thus remain effectively "dominant", even when its absolute proportion of the population is well below an absolute majority.

4   Enzymatic Coupling (Parasitism)

We now define a more general replicase binding rule that will permit enzymatic coupling between different molecular species. As a simple generalisation for investigation, we specify that a replicase will successfully bind (and thus cause replication) as long as it is a *substring* of the template. (We still allow identity binding—i.e., a string binding to another instance of itself—as a special case of "substring" binding.)

As soon as we introduce any such enzymatic coupling between species we necessarily introduce the possibility of exploitation or parasitism: where one species may help replicate another without receiving any reciprocal support for its own replication.[6] Or equivalently, we might refer to such species as, respectively, *altruists* and *egoists*.

Let us, for the time being, discount mutation again. Let the population be initially dominated by a species X (concentration 1.0). Let us introduce a small quantity of another species Y, where the sequence of X is a substring of that of Y. The interactions are then as shown in Table 2.

The corresponding approximate differential equation becomes:

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

Since $x, y \in [0, 1]$ this means that $\dot{x} < 0$ for any non-zero y. That is, such a species Y can invade and completely displace X even from (in the continuous approximation) arbitrarily small initial concentrations. Note the contrast here with the earlier simple "survival of the common" outcome, characteristic of hyperbolic growth. Now the initial dominance of X is of no protection at all against selective displacement. This is because, in this scenario, a high concentration of X delivers just as much benefit to Y as to itself; whereas, Y delivers benefit *only* to itself.

---

[6] Note that in this particular model we can have only *facultative* as opposed to *obligate* parasitism. That is, any parasite will also, by definition, be able to catalyse its own replication even in the absence of its putative host.

5   Mutation + Parasitism

If we now take the system with substring-based enzymatic coupling, and re-introduce replication error (molecular mutation) we encounter a phenomenon which, although fundamentally still Darwinian, seems at least a little counter-intuitive.

Briefly, we suppose that the replication errors include bit-flips, deletions and insertions (through bit repeats). Given reasonable choices of parameters, and initialising the population with a single homogenous sequence, the dynamic will be as follows. In the first instance, the concentration of the initial sequence will decrease in accordance with the analysis of section 3; but it will remain dominant. However, in due course, a mutant will arise which is one bit longer, but otherwise identical to the initial sequence. This will behave exactly as a parasite; and will displace the previous dominant sequence in accordance with the behaviour described in section 4, and become a new dominant species. But, of course, this process will then just be repeated with another mutant, one further bit longer.

Figure 3 shows an example experimental run, illustrating that this is, indeed, precisely the way the system behaves.
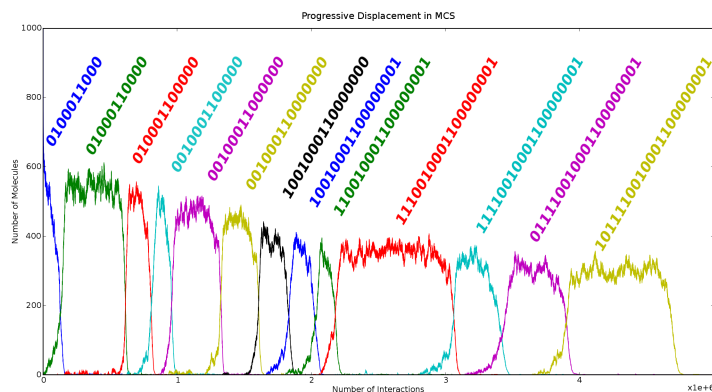

Fig. 3. Displacement Events in MCS

Note carefully that, over this macro-evolutionary epoch, the length of the dominant sequence grows steadily.[7] This means that the per-molecule mutation

---

[7] We note that this property of parasites being more "complicated" than their hosts—at least in the sense of being longer—seems to contrast with the more typical situation with biological parasites, which may be radically simpler than their hosts. This choice is motivated in the toy model presented here simply by the objective of demonstrating the "in principle" *possibility* of sustained macro-evolutionary deterioration in intrinsic fitness.

rate is steadily increasing; or, equivalently, the replication "fidelity" is getting steadily lower. That is, this macro-evolutionary trajectory actually results in a progressive and systematic *deterioration* in intrinsic fitness of the dominating species. This is in marked contrast to the naive "hill climbing" interpretation of evolution; and illustrates how evolutionary processes may be much more a matter of ecological interaction, or game playing, than any kind of optimisation. In particular, we note that this behaviour is completely at variance with the replicator world scenario which Dawkins has previously characterised (or, unintentionally, caricatured) with the slogan "fidelity, fecundity, longevity" [1].

6   Protocells: Two Interacting Levels of Selection

We now extend the model by confining the "template-replicator world" of the previous sections into a "protocell" container. The overall rules of the system are now changed as follows. When a replication event occurs, a molecule is no longer removed. Instead, the absolute number of molecules in the protocell is allowed to increase. However, at a specified maximum size, the protocell spontaneously undergoes binary fission to give two offspring protocells, each containing a random selection of half of the molecules from the parent. There is now a finite limit on the number of protocells. Thus, when a protocell reproduces, another protocell is chosen at random and discarded (along with its entire molecular contents).

A "mutation" at the protocell level now corresponds to the emergence of a molecular parasite as described in section 5. In this simple model, there is no exchange of molecules between protocells except through descent. Accordingly, a molecular parasite will give rise to a lineage of protocells which will necessarily become dominated by the parasitic molecular species. *However:* once this lineage "matures" in this way, the steady-state concentration of the dominant sequence will be slightly lower than in the parental lineage, in accordance with section 3 above (as the per-molecule mutation rate is slightly higher). This means, in turn, that the absolute rate of molecular replication is slightly lower (since this is essentially determined by the square of the concentration of the dominant molecular species). This means that, *at the protocell level*, this new protocell-mutant lineage will be effectively selected against. Thus, protocell level selection will prevent the progressive trend toward longer molecules and lower molecular replication fidelity, described in section 5. Instead, while such protocell mutants will occasionally arise, on an ongoing basis, they will be consistently eliminated again.

Figure 4 shows a test case to illustrate the robustness of this claim. The protocell population is initialised with equal numbers of two cell types. One is dominated by molecules 10 bits in length, the other by molecules of 11 bits in length, where the former is a substring of the latter. However, because of the isolation provided by the protocells, the shorter molecules are now effectively protected against exploitation; and because of their slightly lower replication error rate, and correspondingly higher steady state concentration, the lineage of

protocells dominated by these molecules is able to reliably displace the lineage dominated by the longer molecules.
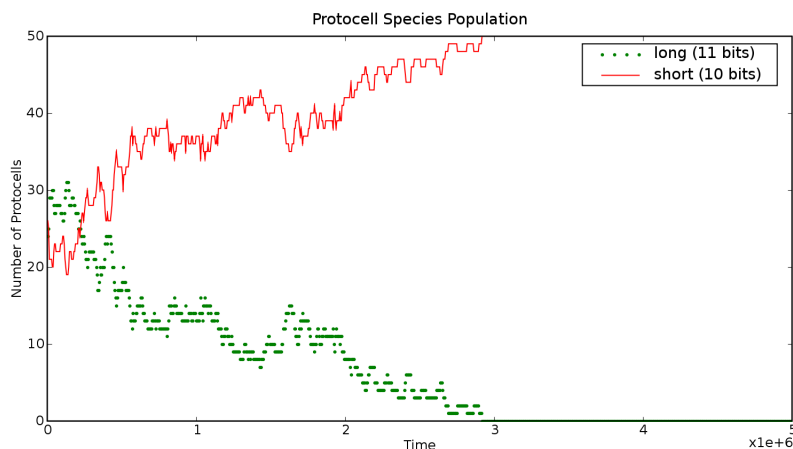

Fig. 4. Protocell Population Stability

It follows that, in theory, at the protocell level, the population could still be taken over by a new lineage of protocells which are dominated by molecules which are shorter again. However, this cannot happen in this model due to *selection at the molecular level*. That is, although, on an ongoing basis, shorter substring molecules will arise in individual protocells, they will be unable to build up any significant concentration relative to their already-dominant parasites (per section 4 above) or even just to any unrelated dominant sequence (per section 2). Thus, because of the molecular level selection dynamics, it is actually not possible to create a mutant *protocell* that is dominated by a shorter length molecular species.

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 that 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.

7   Conclusion: Implications of the Study

The system presented here is, of course, radically simplified compared to any phenomena of real chemistry or biology. Its purpose is not to directly model

such real systems. Rather it is presented as a deliberately minimal system which already illustrates how complex and counter-intuitive the evolutionary behaviour of such systems can be; but also, how the evolution can, indeed, be dramatically altered by the interaction between multiple levels of selection.

The broader intention of the 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. We therefore conjecture that protocellular computation is precisely a phenomenon that emerges in the conditions of a "major transition" between evolutionary levels. Further work will involve incrementally widening the repertoire of molecular interactions while working to still understand how this impacts on both interacting levels of evolution.

Acknowledgements

References

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

[2] Peter Dittrich, Jens Ziegler, and Wolfgang Banzhaf. Artificial Chemistries - A Review. *Artificial Life*, 7(3):225–275, 2001.

[3] Manfred Eigen and Peter Schuster. The Hypercycle, A Principle Of Natural Self-Organization. *Die Naturwissenschaften*, 64(11):541–565, 1977.

[4] Walter Gilbert. Origin of life: The RNA world. *Nature*, 319(6055):618, February 1986.

[5] GF Joyce. The rise and fall of the RNA world. *The New Biologist*, 3(4):399–407, 1991.

[6] John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. Oxford Press, 1997.

[7] Barry McMullin. Replicators don't. In Federico Morán, Alvaro Moreno, J. J. Merelo, and Pablo Chacón, editors, *Advances in Artificial Life, Third European Conference on Artificial Life, Granada, Spain, June 4-6, 1995, Proceedings*, volume 929 of *Lecture Notes in Computer Science*, pages 158–169. Springer, 1995. `http://www.eeng.dcu.ie/~alife/ecal95/rpl-l2h/rpl-l2h.html`.

[8] Eörs Szathmáry and John Maynard Smith. From Replicators to Reproducers: the First Major Transitions Leading to Life. *Journal of Theoretical Biology*, 187:555–571, 1997.