# Macroscopic Structures Emerging from the Interactions of Simple Replicators

Perry W. Swanborough

## Abstract

A macroscopic view of Langton (1984) and Byl (1989) self-replicating loops has facilitated their representation by oriented states of single cells. In this work, the single-cell representation of the loop replicator is adapted to facilitate a wide variety of interactions identifiable by observation of the dynamics of arbitrary spatial distributions of many replicator copies. Dynamic loops of periods greater than the four-step rotation cycle of the orientated state were observed, and examples with periods of 20 and 12 are shown. The paper includes some discussion of irreversibility in both computation and biology, with recognition of the problem of deriving concrete conclusions about abiogenesis and pre-LUCA biology from speculative preliminary abstraction.

*Keywords*: abiogenesis, artificial life, cellular automata, entropy, information, irreversibility, loop replicator

## Introduction

There have been several efforts to deduce features of the Last Universal Common Ancestor (LUCA) of current life from the observed universal features of extant life. Though conclusions across LUCA studies differ, a comparison of eight studies [3] revealed a consensus that the LUCA incorporated a genome which facilitated amino acid and nucleotide metabolisms supporting protein synthesis, indicating a high complexity baseline. The unavoidable recognition of complexity of LUCA indicates much developmental distance between abiogenesis events and LUCA [2] so it is rational to assume that features of pre-LUCA life were different from the features of LUCA. Therefore, investigation of possibilities for pre-LUCA life requires facing the challenge of avoiding features of extant life as much as possible while not excluding the potential for developmental pathways toward the features of contemporary extant life via LUCA.

The historical transitions from pre-biotic conditions to ancestral replicators/metabolism/cells to subsequent lifeforms are of ongoing interest to the research community, and study of simple abstract replicating structures may be of use in recognizing any universal principles. Among many computational exercises inspired by life, Christopher Langton constructed a self-replicating cellular automata (CA) structure (a self-replicating loop) which replicates in 151 cell state transition steps [7]. The detailed loops can be described more simply as orientated states of single cells to provide a macroscopic view of replication dynamics over many cycles of replication within an extended spatial range [7][8]. Before introducing the single-cell replicator abstraction, replication of the detailed loop is summarized:

### Langton's self-replicating loop

An instance of Langton's loop [7] replicates into quiescent (state 0) space in the direction pointed at by its construction arm. When a descendant loop is completed, replication of the parent is redirected 90 degrees counterclockwise enabling a subsequent replication in the new direction, subject to available quiescent space. If and when replication of a loop is blocked by the presence of an already-established structure, erasure of the internal instruction tape initiates, and on completion of

information erasure only a static skeleton of the previously-active loop persists. Figure 1 shows a colony of loops developed from an initial loop in isolation.

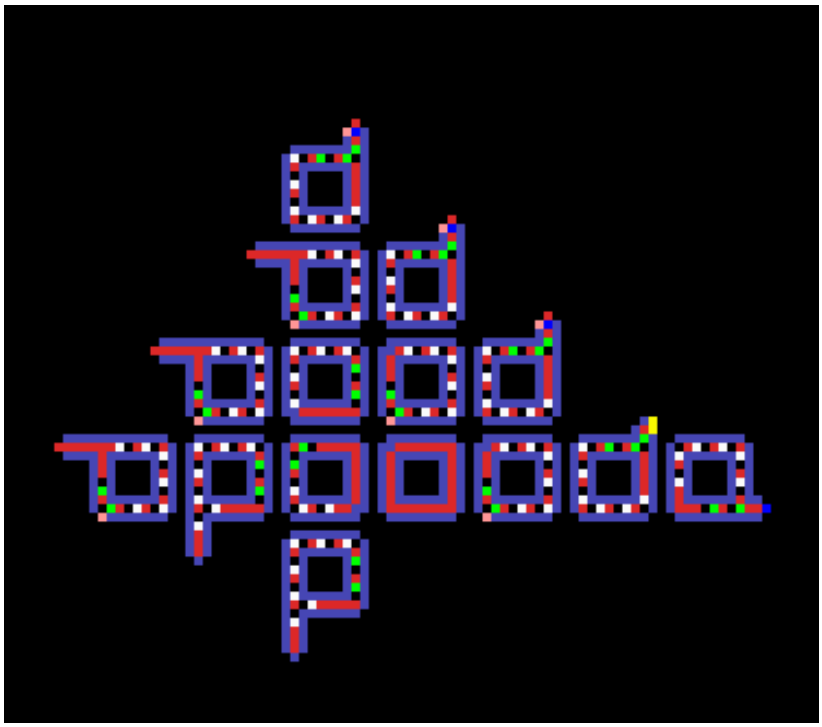


**Figure 1**. The result of 595 iterations from one isolated Langton replicator [7] at Time = 0. Cell colours to states correspondences: black = 0 (quiescent), red = 1, deep blue = 2, pink = 3, green = 4, yellow = 5, bright blue = 6, white = 7. The 595 iterations correspond to four replications of the ancestral structure (third from the left in the horizontal row of seven), with replications of most of the outer boundary structures ongoing. The Figure was produced using *Processing* [9].

In Figure 1, the initial replicator loop is the third from the left in the horizontal row of seven. At the 595$^{th}$ iteration shown, partial erasure of the internal information tape has occurred and continues to completion with subsequent iterations. Note that the initial replicator's first descendant is to its immediate right (East), and at the 595$^{th}$ iteration erasure of this descendant's information tape is already complete - it completed only two replications before information erasure was triggered by blockage of replication to its left (West) by its own parent.

**Replication of Byl's simplification of Langton's loop**

Given the intent of investigating how compact a non-trivially replicating structure could be, John Byl derived a simplification of Langton's replicator [1]. A colony of instances of Byl's replicator develops from a single isolated ancestor similarly, with the exception that no information erasure occurs in loops spatially blocked from an otherwise possible replication (Figure 2). The internal information tape (the 2x2 arrangement of cell states 1, 3, 3, 4) of spatially blocked loops cycles unproductively and indefinitely.

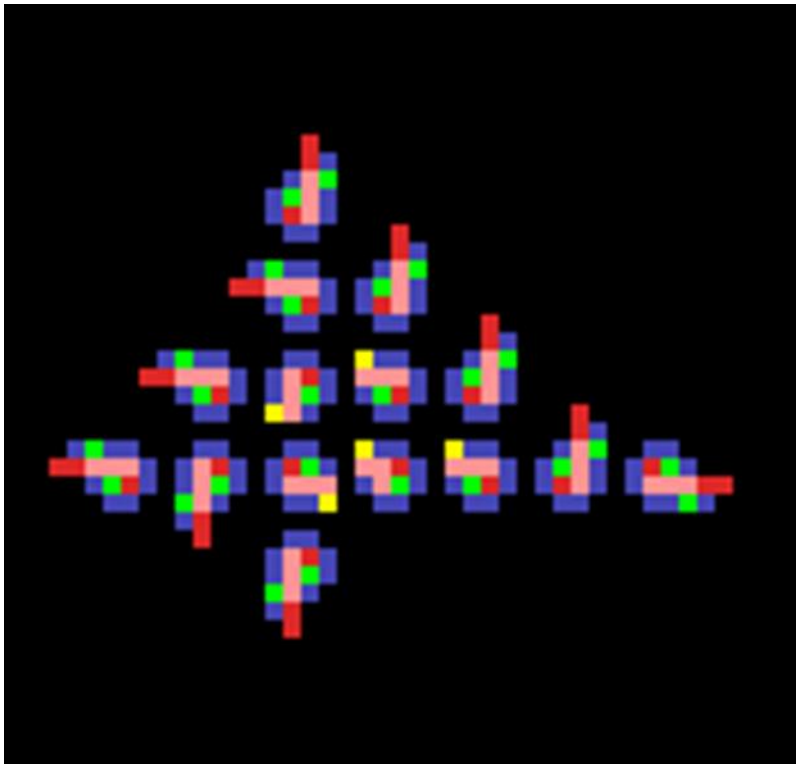


**Figure 2**. The result of 105 iterations from one isolated Byl replicator at Time = 0. Cell colours to states correspondences: black = 0 (quiescent), red = 1, blue = 2, pink = 3, green = 4, yellow = 5. The 105 iterations correspond to four replications of the ancestral structure, with replications of most of the boundary structures ongoing. The Figure was produced using *Processing* [9].

Additional to describing the structure and replication of the loops summarized above, a macroscopic view of loop replication, *i.e.* loops abstracted to orientated states of single cells, was implied by Langton (Figure 10 in [7]) – but the macroscopic view was not discussed beyond the context of the rigid colony development from a single initial replicator. Mange *et al.* also later reviewed the macroscopic view of Langton's replicator, specifically for their development of programmable digital logic applications of self-replicating structures [8].

As a prospective tool for thinking about abiogenesis and protobiology, what can the study of simple abstract replicating structures tell us? Loop replicators are brittle, *i.e.* not robust against variation and accordingly, development from a single instance only generates the rigid colony form illustrated in Figures 1 and 2. Accepting the constraints implicit in loop replicators as primitive units, how can a variety of larger scale/higher order structures and functions develop from the interactions of randomly-distributed multiple replicator copies? In this work, the macroscopic view of loop replicators [7][8] is further developed to investigate a wider range of dynamics within a spatial distribution of loop replicators, with the objective of identifying any emergence over larger scales of space and time.

## Methods

To investigate large-scale dynamics of arbitrary distributions of interacting replicators, the abstraction of loop replicators as orientated states of single cells is revisited. These oriented single-cell replicators inherit both the spatially-directed replication of Langton and Byl loops, and the rotating direction of replication, but the abstraction enables facilitation of features expected of dynamic biologically-relevant systems which are difficult or perhaps not possible at all to implement at the level of detail of the original replicating loops. Figure 3 illustrates replication and rotation of a loop replicator represented by the oriented cell state **>**.

**>0 → ^>**

**Figure 3**. Replication from Time t (left) to t+1. A replicator replicates into quiescent space (state **0**) and rotates counterclockwise for further replications into available quiescent space (see [7] and [8]).

### Additional rules

In the development of colonies from one isolated ancestral replicating loop [1][7], replication was organized to be permanently disabled when otherwise-possible replication is obstructed by an already existing loop. An alternative dynamic implemented in this work is to allow continuing rotation of prospective replication so that subsequent successful replication can occur into quiescent space where available. Under this dynamic, replication ceases only when the replicator is surrounded by neighbouring structures in all N, S, E and W directions. By substitution of oriented states for replicator loops, and allowing this new feature, a colony developing from a single isolated ancestor is shown at four state transitions in Figure 4.

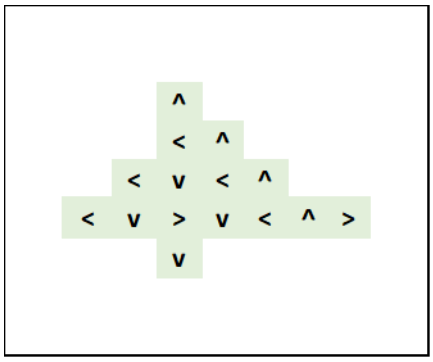



**Figure 4.** A single replication of Langton or Byl loop replicators corresponds to just one time increment in the system of corresponding single-cell, oriented state replicators. The figure shows a colony of replicators at four state transitions from an isolated ancestor **>** (compare with Figures 1 and 2). White space corresponds to quiescent cells (state **0**).

An alternative to the persistence of a replicator skeleton [7] or a blocked replicator unproductively rotating [1] after replication is disabled is implementation of decay, *i.e.* transition of a fully-blocked replicator structure to the quiescent state 0, which can be thought of as death by overcrowding. In Figure 4, three of the shown active states will revert to state 0 at the next iteration under this rule.

Note that under this rule there is no distinction between “dying” and “dead” loops recognized in [7] and [8].

With a generalized random or arbitrary spatial distribution of many replicators, the potential problem of more than one replicator pointing to a single quiescent cell for a prospective replication presents. With replicators being identical, there is no criterion for favouring one replication into a neighbouring quiescent cell over the replication of coexisting others prospectively able to replicate into the quiescent space. The unbiased solution is to exclude any replication at all where there is more than one otherwise-possible replication into a single quiescent cell. The solution is illustrated with a case of two replicator instances in Figure 5 below.

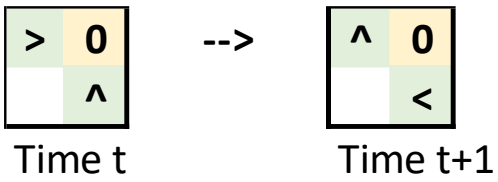



**Figure 5.** At Time t, two replicator instances are pointing to one quiescent cell (state 0). Either instance in isolation would replicate a copy into the quiescent cell (as in Figure 3), but with no possible unbiased criterion for favouring replication by any particular instance, no replication at all occurs. Note that the normal counter-clockwise rotation of all replicators is unaffected.

A further feature of biological systems not implicit in the development of loop replicator colonies is a death mechanism restricting the indefinite expansion of the colonies. The prospective new mechanism is that just as replication expands into quiescent space, quiescent space can intrude onto neighbouring replicator states. The mechanism is additional to the death by overcrowding rule already described above. While excluding immigration and emigration dynamics for simplicity, ideal biological populations persist under an average balance of opposed death and reproductive forces.

Introducing the intrusion of quiescence effect comes with the problem of achieving quantitative balance against replication growth. Each oriented replicator state replicates no more than once each time increment, but because the quiescent state is not oriented, there can be no arbitrary bias in the direction it expands. To achieve the quantitative balance required, four successive time increments are permitted to allow up to four replications by each extant replicator state. The fourth time increment incorporates both the intrusion of each quiescent state 0 onto all (up to four) immediate neighbour replicator states and the usual replication/rotation dynamics of the surviving replicators. (I concede that this is somewhat contrived.) A repeating four-step sequence of three consecutive single-phase transitions of state replication and rotation interspersed with one two-phase transition incorporating both quiescence intrusion and state replication/rotation drives the development of a spatially-extended distribution of replicators over time.

The two-phase state transitions include rules which contradict rules in the single-phase state transitions, as the intrusion of quiescence component of a two-phase state transition erases replicator cells which would otherwise rotate and persist under the rules of a single-phase state transition. Clearly, pooling of rules of both single-phase and two-phase state transitions into one comprehensive state transition function corresponds to a global state transition function containing contradictory state transition rules.

Recognizing a single global state transition function as unviable leaves us with an alternative: if we describe a set of state transition rules specific to a Time = t to t+1 state transition as a state transition *operator*, a state transition function can be redefined as a comprehensive set of internally-consistent state transition *operators*, each operator specific to a respective Time t to t+1 state transition. The absence of rule contradictions within each of the state transition operators holds for all of the single- and two-phase single step state transition operators applied in Figure 6.

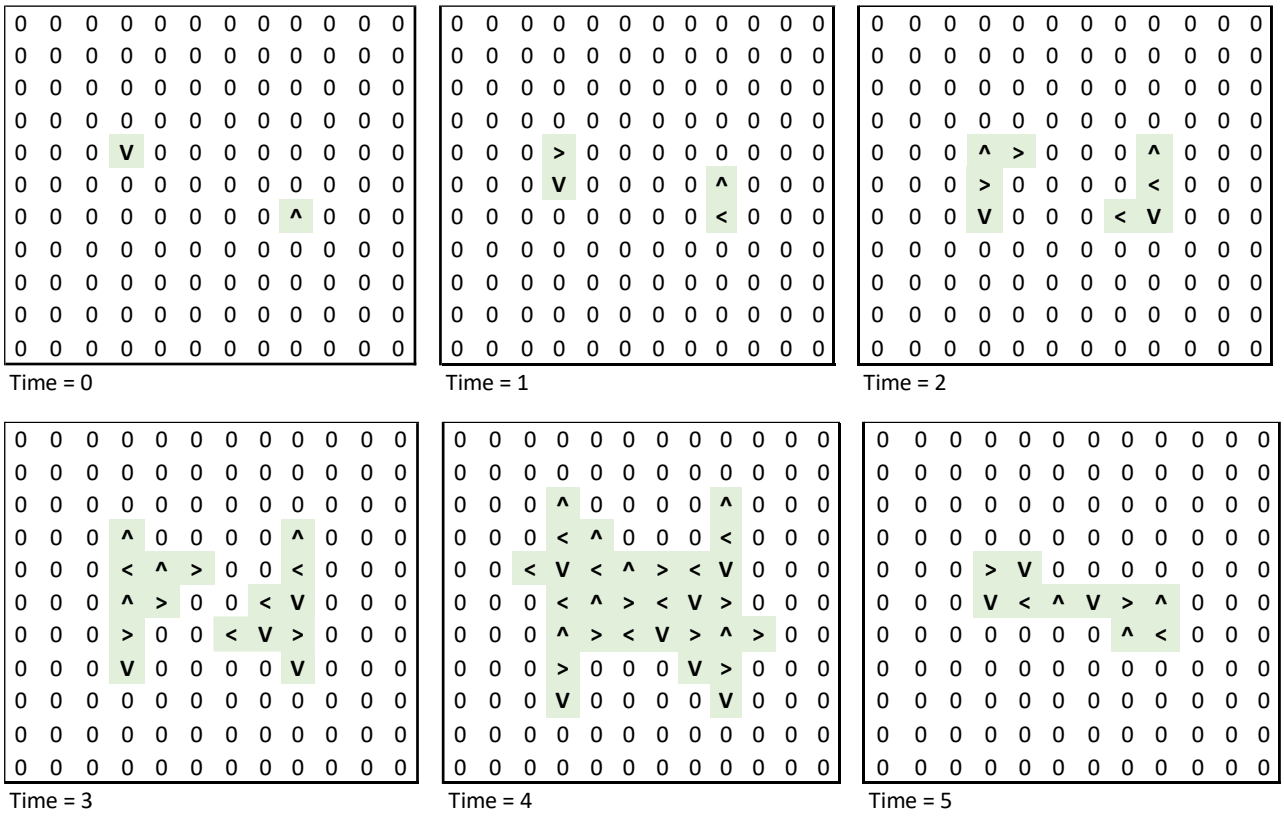
**Application of these rules**

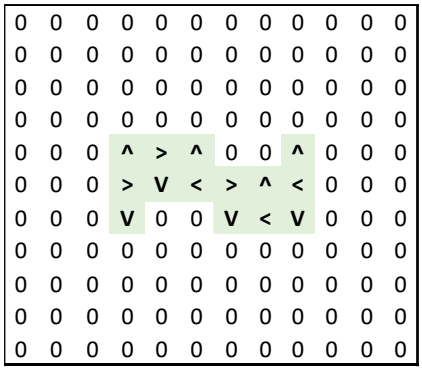
States **^**, **<**, **v**, **>** and **0** were randomly allocated to a 720 x 720 cell array with equal probabilities, determining the initial active state expectation density at 0.8. The array was implemented with toroidal topology to exclude spatial edge complications. Dynamic behaviour within the grid was observed as the state transition rules were applied.

## Results

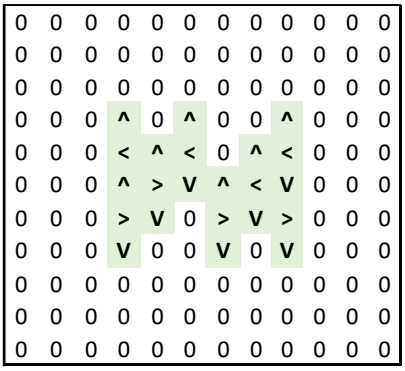
From the interaction of many replicator instances, this exercise revealed an expanded range of dynamics in a CA environment equipped with the rules described in Methods.

A general four-step cycle over a spatially-extended range was observed as active states rotated through the four orientations, but 12- and 20-step dynamic cycles developing from simple isolated origins also occur. A 20-step cycle developing from just two spatially-separated, isolated replicators at Time = 0 is illustrated in Figure 6 below.

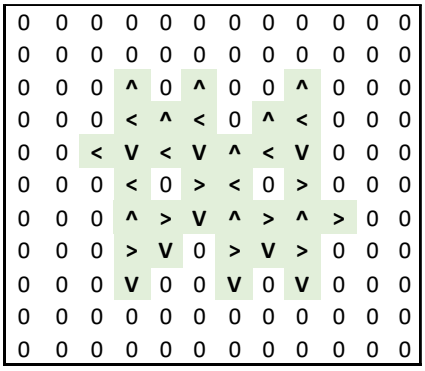




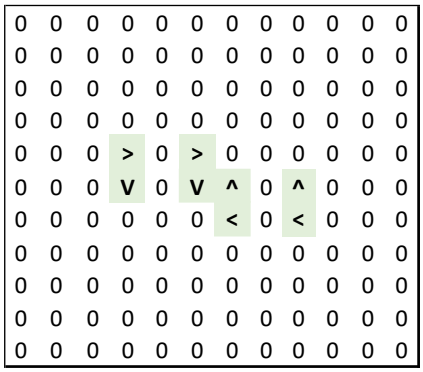
Time = 6



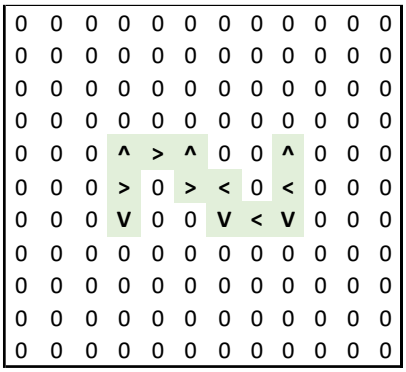
Time = 7



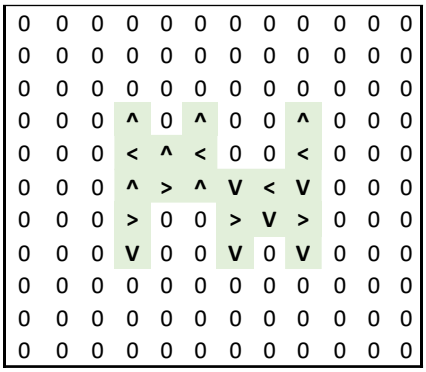
Time = 8



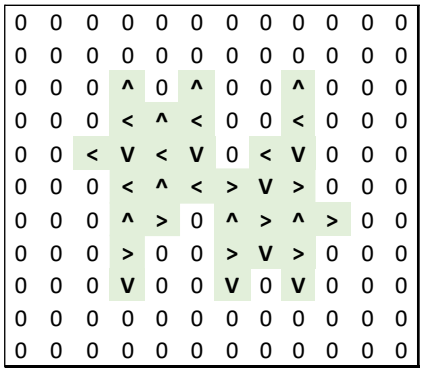
Time = 9



Time = 10

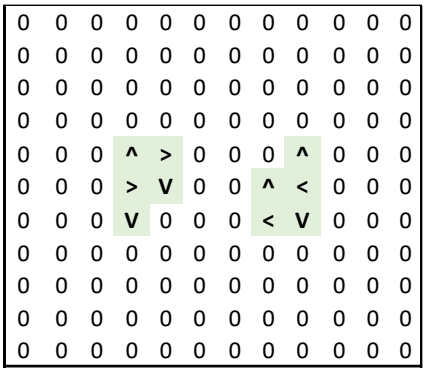


Time = 11

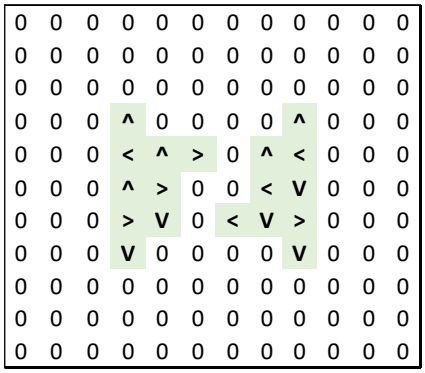


Time = 12

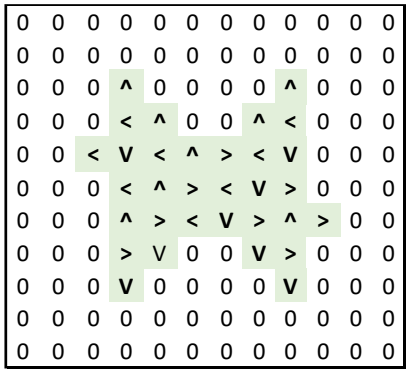
Time = 13



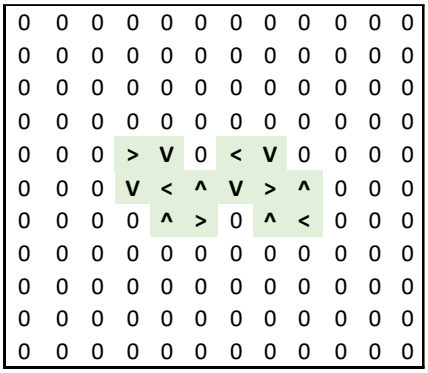
Time = 14



Time = 15



Time = 16



Time = 17

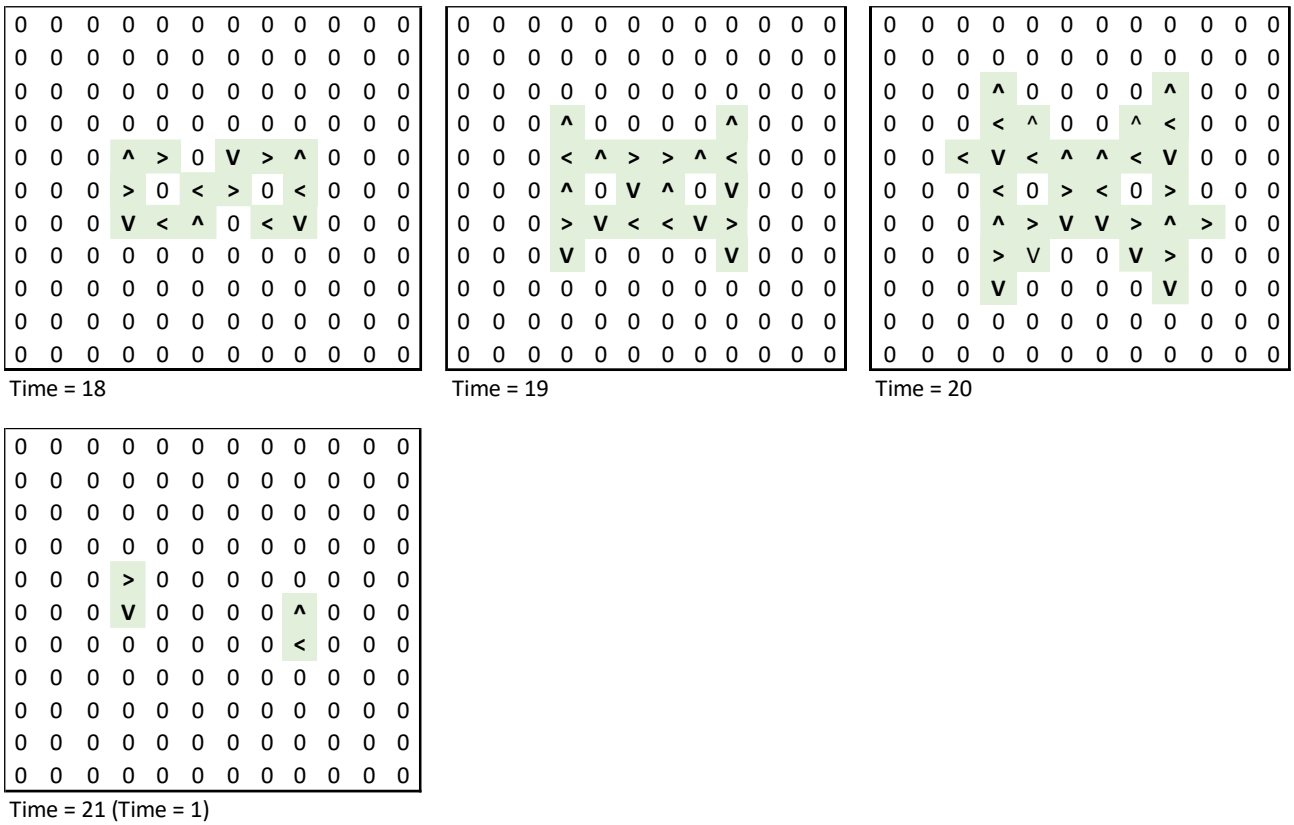



**Figure 6**. A cycle of period 20 driven by a sequence of applied single- and two-phase state transition operators (see text). Two-phase state transition operators are applied every fourth state transition from Time = 1 (*i.e.* from Times 4 to 5, 8 to 9, 12 to 13, 16 to 17, and 20 to 21). Figure 7 below shows detail of a two-phase state transition example.

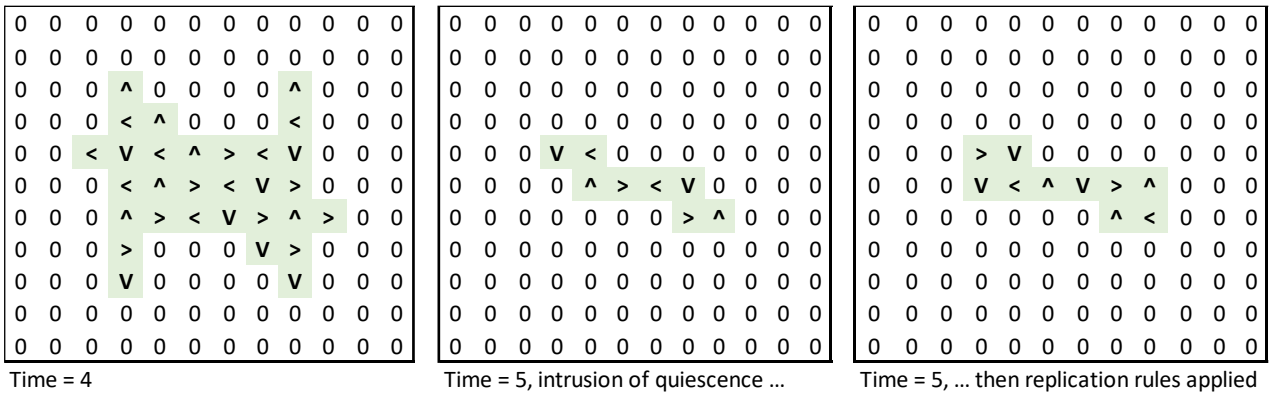



**Figure 7**. The Time = 4 to 5 state transition shown in Figure 6 exemplifies a two-phase state transition operator. The single time-step, two-phase state transitions incorporate a reduction of the population of replicators by the intrusion of quiescence (left frame to middle frame), then a normal replication and rotation of the surviving replicators (middle to right frame).

An Appendix at the end of the paper details a cycle of twelve steps (Figure 8).

## Discussion

A CA environment of structures and state transition operators, as exemplified in Figures 6 and 8, constitutes a dynamic system in which these two categories mutually interact: each structure corresponds to a set of CNESW neighbourhood states which match the rule-content of a corresponding state transition operator. The operator so selected by the structure then modifies the structure, and a cycle of operator selection and structure modification perpetuates. As products of historical dynamic interactions, the structures can potentially serve as repositories of memory, but while memory can accumulate it is also subject to deletion, with obvious limitations to study of abiogenesis and protobiology from unavoidably-limited extant evidence. Memory deletion implies irreversibility, as discussed below.

### Irreversibility in biology

Logical degeneracy of the translational mapping of RNA codons to amino acids facilitated by ribosomes (*i.e.* the genetic code) implies irreversibility of RNA-to-protein translation. Specifically, 61 codons correspond to 20 amino acids so prospective inverse translation is many-valued (only two amino acids correspond to a unique codon each), though any prospective inverse translation in principle would produce a codon sequence functionally equivalent to the specific sequence initially generating the protein. Such a reverse translation would require a mechanism of arbitrary or stochastic mapping of 18 of biology's 20 amino acids to one particular corresponding codon each.

The Central Dogma of Molecular Biology is recognition that information encoded as polypeptide amino acid sequences is not reverse-translated to nucleic acids, and we might assume that the reason is degeneracy of the genetic code. However, the Central Dogma appears to hold for reasons other than genetic code redundancy. The reason for the Central Dogma was recently argued to be a fundamental impossibility of deriving linear RNA codon sequences from the spatially-complicated 3D tertiary structures of corresponding polypeptides [5]. The denaturing of 3D polypeptides/proteins to fully-intact readable 1D amino acid sequences is for all practical purposes thermodynamically impossible.

### Irreversibility of loop replicator dynamics

In each forward (Time = t to t+1) CA state transition, all instances of any one CNESW neighbourhood state at Time t correspond to one C to C' state transition at Time = t+1, *i.e.* there are no rule contradictions within any of the forward-direction state transition operators. However, considering corresponding state transitions in the reverse Time = t+1 to t direction, there are prospective state transition operators where more than one C state at Time = t corresponds to a state C' at Time = t+1 over multiple instances of one Time = t+1 CNESW neighbourhood, *i.e.* there are forward state transition operators with multi-valued inverses. These forward state transition operators are logically irreversible.

**Examples: irreversibility of some single-phase state transition operators applying in Figure 6**

By inspection of the forward-transition from Time = 7 to 8, the C = **>** state within the CNESW neighbourhood **>^vv^** at Time = 7 transitions to C' = **0** in neighbourhood **0<>><** at Time = 8. Simultaneously, the C = **<** state within the CNESW neighbourhood **<^vv^** at Time = 7 transitions to C' = **0** in neighbourhood **0<>><** at Time = 8. We can immediately see in the reverse transition Time = 8 to 7 that state C' = **0** at Time = 8 in each of the instances of the CNESW neighbourhood **0<>><** transitions to state C = **>** in one of the instances and to C = **<** in the other. The multi (two)-valued inverse state transition rule corresponds to irreversibility of the forward state transition from Time = 7 to 8.

Another example is apparent by inspection of the forward-transition from Time = 17 to 18. The C = **<** state within the CNESW neighbourhood **<v^^v** at Time = 17 transitions to C' = **0** in neighbourhood **0><<>** at Time = 18. Simultaneously, the C = **>** state within the CNESW neighbourhood **>v^^v** at Time = 17 transitions to C' = **0** in neighbourhood **0><<>** at Time = 18. We can immediately see in the reverse transition Time = 18 to 17 that state C' = **0** at Time = 18 in each of the instances of the CNESW neighbourhood **0><<>** transitions to state C = **<** in one of the instances and to C = **>** in the other. Again, the two-valued inverse state transition rule corresponds to irreversibility of the forward state transition from Time = 17 to 18.

Note that the irreversibility of these single-phase state transition operators is due to the reversion to the quiescent state of active replicators fully-blocked by replicator neighbours at all N, S, E and W directions, *i.e.* erasure of prior active states corresponds to irreversibility.

**Irreversibility of two-phase state transition operators applying in Figure 6**

Fourteen of the state transitions in the Figure 6 loop are reversible, but as shown above, the Time = 7 to 8, and 17 to 18 single-phase state transitions are irreversible. Additionally, all of the two-phase state transitions are irreversible due to erasure of prior replicator states by intrusion of quiescence, as detailed below:

Reverse-mappings of C' = state **0** within neighbourhood **0**0000 at Time = t+1 onto any of C = **^**, **<**, **v**, or **>** at Time = t are equivalent under rule rotational symmetry, but any of these state-transition rule rotations coexisting within one reverse state transition operator mutually contradict each other. Even considered as one rule in recognition of rotational symmetry, it contradicts the rule of state C' = **0** within neighbourhood **0**0000 mapping to state C = **0** applying in uniform quiescent space. We see that irreversibility is an outcome of the intrusion of quiescence.

Irreversibility due to the intrusion of quiescence can also be seen in the Time = 4 to 5 two-phase state transition shown in Figures 6 and 7, which includes CNESW → C' transitions **^00<<** → **0** and **>0<<^** → **0**. By inspection of the corresponding reverse transitions, the prospective Time = 5 to 4 state transition rules are **000v0** → **^** and **000v0** → **>**. The contradicting C values of **^** and **>** corresponding to a reverse state transition from C' = **0** within instances of the CNESW neighbourhood **0**00v0 at Time = 5 demonstrate irreversibility.

Note that irreversibility does not prevent the reappearance of past macro-states, *i.e.* loops as exemplified in Figures 6 and 8 occur.

**Irreversibility of physical (and computational) systems**

Traditionally, information processing has been considered the function of organic nervous systems, and our artificial computing hardware, but increasingly, all processes in biology are being considered as manifestations of information processing [4]. Concepts of thermodynamics, irreversibility and entropy discussed in the context of information processing may therefore be fruitful in extension to discussions of biological processes.

The erasure of information corresponds to a cost of increased entropy, recognizable as the dissipation of heat [6]. The increased entropy indicates the irreversibility of a dynamic step when information implicit in prior causal states is lost. In more abstract terms, any forward dynamic step for which the corresponding inverse is multi-valued is irreversible when a record of the singular preceding cause is erased. It follows that if nothing is erased as computation proceeds then a procedure is reversible in principle, but at a cost of perpetuating an associated memory of indefinite size sufficient to hold details of all prior causal steps.

Entropy generated as CA computations proceed presumably corresponds to some portion of the heat dissipated by the hardware running the CA software, but the total heat dissipated by computing hardware includes entropy generation by electrical resistance and computational overhead not directly attributable to the CA computations. Empirical quantification of entropy generated directly by CA dynamics is not a trivial problem. Does it make sense to discern information-processing entropy (dissipated heat) from the balance of total dissipated heat, given that the information processing only occurs by all of the operations of the hardware with the associated total of entropy generation? The question of separability of information-processing heat from other overhead heat generation was articulated well and discussed in [10].

**The problem of deriving the concrete from the abstract.**

Whatever insights may be deduced from studying abstractions of biological phenomena, relation of the abstract to the concrete is essential. Do the abstractions point to specific chemistries and biophysical mechanisms? A comprehensive thermodynamic assessment of CA dynamics potentially relevant to the concrete questions of abiogenesis and proto-biology may need to consider continuity of the software-level CA logic with heat dissipation of hardware to potentially obtain insights more comprehensive than from reasoning limited only to consideration of logical state transitions.

## References


[1] J Byl, Self-reproduction in small cellular automata, *Physica D* **34** (1989) 295-299.

[2] A Cornish-Bowden and ML Cárdenas, Life before LUCA, *Journal of Theoretical Biology* **434** (2017) 68-74.

[3] AJ Crapitto, A Campbell, AJ Harris and AD Goldman, A consensus view of the proteome of the last universal common ancestor, *Ecology and Evolution* DOI: 10.1002/ece3.8930 (2022) 13 pp.

[4] P Davies, *The Demon in the Machine*, Allen Lane, Penguin Random House, UK (2019).

[5] EV Koonin, Why the Central Dogma: on the nature of the great biological exclusion principle. *Biology Direct* (2015; 10:52).

[6] R Landauer, Irreversibility and Heat Generation in the Computing Process, *IBM Journal* (July, 1961) 183-191.

[7] CG Langton, Self -reproduction in cellular automata, *Physica D* **10** (1984) 135-144.

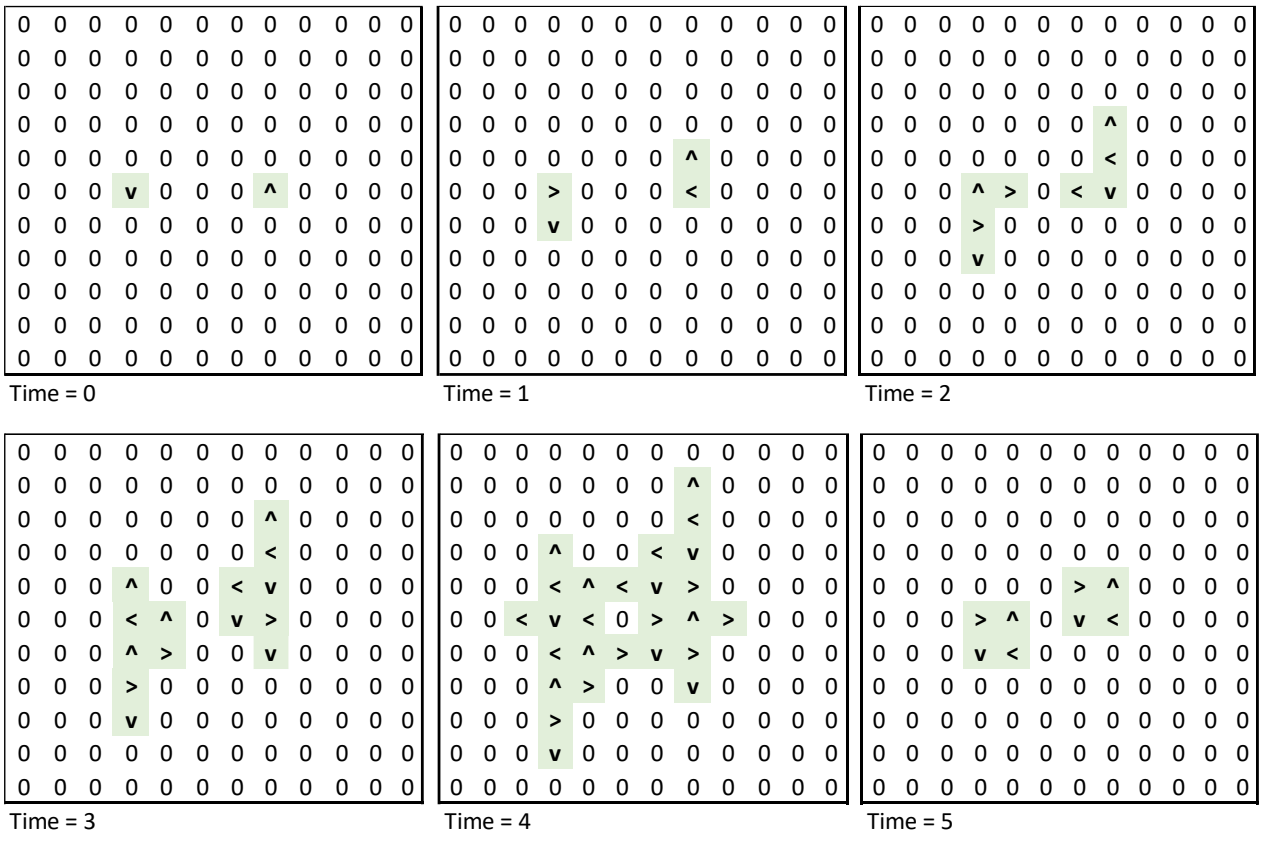
[8] D Mange, A Stauffer, L Peparolo and G Tempesti, A Macroscopic View of Self-Replication, *Proceedings of the IEEE* **92** (2004) 1-17.

[9] C Reas and B Fry, *Processing, A Programming Handbook for Visual Designers and Artists* (second edition). MIT Press, Cambridge MA (2014).

[10] I Sárándi, Why do computers generate heat?, *https://physics.stackexchange.com/questions/137504/why-do-computers-generate-heat* (2014).

## Appendix

Figure 8 below shows a cycle with a period of twelve state transitions.



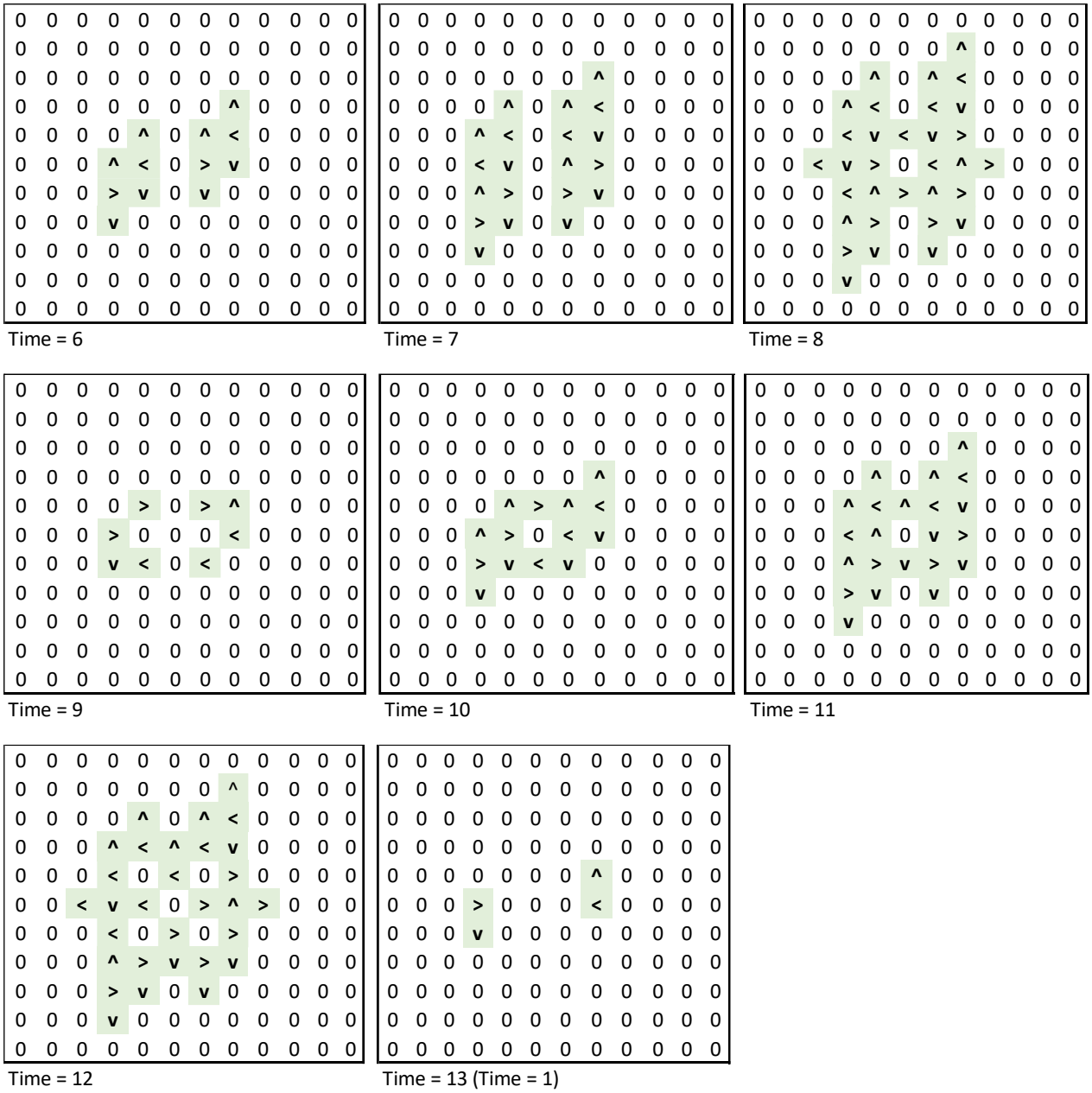



**Figure 8**. A cycle of period 12 driven by an applied sequence of single- and two-phase state transition operators (see text). Two-phase state transition operators are applied every fourth state transition, *i.e.* from Times 4 to 5, 8 to 9, and 12 to 13.