Research Article

On the Increase of Functional Information in Complex Systems

Hank Rainwater[1]

1. Independent researcher

For complex systems of many interacting components, the functional information of a system is expected to increase if many different configurations of the system are subjected to selection for one or more functions according to a recently defined new law. A cellular automata model – a classic example of a complex system – was created as a testbench for this law, and the model's evolutionary results were analyzed and documented by Functional Information theory. In the model, over 800,000 system configurations evolved and created 176 unique functional structures in a first phase of evolution, demonstrating the law's validity. An additional discovery was made that correlates the functional information of a structure with the computational work in evolutionary generations required to create the structure. Sequential phases of evolution of the model build on the results of previous phases and reveal a growth in the complexity of created structures.

Corresponding author: J. Hank Rainwater, rainwater.hank@gmail.com

1. Introduction

New laws of physics are rare in the history of science. The gradual transition from hypothesis to theory along the path of reproducible experimentation is a long journey. A consensus among scientists to promote a theory to a law can take decades, if not centuries, of peer-reviewed research and validation. For self-organizing systems that occur in nature, a current seminal review stated that "After more than a century of concerted effort, physics still lacks basic principles of spontaneous self-organization"[1]. However, a new law was recently proposed, applicable to the roles of function and selection in complex systems described by the following criteria[2]:

1. Each system is formed from numerous interacting units (e.g., nuclear particles, chemical elements, organic molecules, or cells) that result in combinatorially large numbers of possible configurations.

2. In each of these systems, ongoing processes generate large numbers of different configurations.
3. Some configurations, by virtue of their stability or other "competitive" advantage, are more likely to persist owing to selection for function.

The new law for these systems is defined as follows:

The functional information of a system of many interacting components will increase (i.e., the system will evolve) if many different configurations of the system are subjected to selection for one or more functions[2].

The concept of functional information (FI) was first defined by Nobel laureate Jack Szostak in a biological context as "a quantitative means of comparing the functional abilities of different biopolymers would allow us to dissect out differences and to discern their origins"[3]. In later work, FI was used as a theoretical tool to study the emergence of biocomplexity[4] where the theory's applicability was extended beyond its original investigative domain of biology to the study of letter sequences and artificial life, and it has become an important tool for research into the origins of complexity by quantifying complexity by degrees of functionality. The new law of increasing functional information, abbreviated as LIFI, was applied to the evolution of minerals[5] in the context of open-ended versus bounded evolution and offered as one example, among others, of increasing functional information in a recent book written for a general audience by the same two scientists[6].

In previous research on artificial life in cellular automata (CA)[7], I demonstrated how a modified version of Conway's Game of Life (GoL)[8], called perennial evolution (PE), can create stable and increasingly complex structures characteristic of complex systems of many interacting parts with the ability to "dissipate energy or to maximize entropy production through patterning"[4]. PE is identified as *driven* because it continually injects order during evolution by keeping a single CA cell or a previously constructed structure permanently active and extends the system's dissipative action due to the physics of its state transition rule (STR)[9]. The structures created in PE – hereafter referred to as patterns – display a selection bias toward certain geometric forms whose creation is a result of chemistry-like reactivity with the embedded, permanently active cell(s) acting as a catalyst, functioning as an artificial chemistry. PE satisfies the criteria of a system that could demonstrate an increase in functional information and will be used in this research to test the applicability of LIFI as an analysis tool.

2. Perennial Evolution Design

Pattern construction in the first phase of evolution begins with two types of CA cells, as shown in Figure 1.

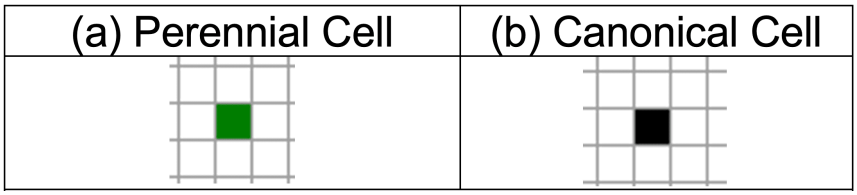


Figure 1. Construction Components. Building blocks for construction of patterns in PE.

A single perennial cell exists amid a collection of randomly positioned canonical cells in a toroidal CA grid called a *soup* in the GoL parlance. The pattern-forming interaction between a soup's cells is controlled by the physics of the STR acting in discrete steps (*generations of evolution*) on all cells synchronously in the grid (*universe*) according to the following rules:

1. Each inactive cell is defined as having the potential to become an active canonical cell and will become active if it has 3 active neighbors; otherwise, it remains inactive.
2. A canonical cell that is active remains active during evolution if it has 2 or 3 active neighbors; otherwise, it becomes inactive. This rule, combined with rule 1, is often expressed using the nomenclature B3/S23, where 'B' signifies born and 'S' survive.
3. A perennial cell cannot be created by the STR but is manually introduced into the initial configuration of canonical cells as a seed for constructing patterns prior to the beginning of evolution. A perennial cell is permanently active throughout all generations and can have as many active canonical neighbors in the 8 cardinal and intercardinal directions (the Moore neighborhood) of the cell as rules 1 and 2 will allow.
4. A collection of canonical cells bound to a single perennial cell resulting from the action of rules 1, 2, and 3 is considered a functional pattern if it becomes static during construction and remains static until the universe reaches a quiescent end state. This end state occurs when the density of active cells in the universe reaches a constant or periodic value, but the pattern itself must be static. Also, a functional pattern must have bonds connecting all the cells, i.e., it must be a strict still life as defined by the GoL, consisting of only one 'island.'

Many familiar GoL still lifes will be created in PE because rules 1 and 2 are identical to the canonical rules of the GoL and are identified as *canonical* patterns. However, patterns will also be created whose stability is maintained by the single perennial cell, and these are termed *perennial* patterns that cannot exist in the GoL.

The birth of these patterns is a construction resulting from the computation performed by the STR over generations, and the word *evolution* is used not in the biological sense but rather as a shorthand term describing a complex system whose behavior is analogous to self-organizing processes in nature driven by the physics of the interaction of the units that constitute the system. While the stable structures formed by PE can be considered artificial life or artificial molecules, the term pattern will be used to clarify that it is a model *simulating* a natural process. In a Qeios preprint article, I defined this model's phenomenology in detail, but this article focuses on the *increase* in functional information displayed by the model. (The software developed for the model is available for download via GitHub).

3. Functional Information Defined for Perennial Evolution

FI theory defines a "degree of function (E_X) as a quantitative measure of a configuration's ability to perform the function x"[10]. This *function* 'x' in PE is the ability of the perennial cell to create a stable pattern during the evolution of a random soup, where an unevolved soup is defined as a configuration of the system. The "quantitative measure" of the "degree of function" for a pattern is given by the formula:

$$\mathrm{FI}(\mathrm{E}_X) = -\log_2[\mathrm{F}(\mathrm{E}_X)] \qquad \text{(Equation 1)}$$

In Equation 1, $F(E_X)$ is the "fraction of all possible configurations of the system that achieves a degree of function". In PE, this fraction represents the probability of occurrence of a pattern and is equal to the number of copies of a pattern divided by the total possible number of configurations.

Determining the size of "all possible configurations" for PE is the next step in solving Equation 1. In a study of the changing FI of suites of minerals at different stages of Earth's mineral evolution, a semiquantitative method of calculating the total number of all possible minerals at each era of Earth's geological history was accomplished, exhausting the possibility space[5]. In PE, the possibility space is much larger since configurations in PE are random soups created in a toroidal universe of 31 rows and columns that contain 961 cells with 35%, or 336 cells, that are active prior to evolution. The total possible soups that can be created is a very large number, 5×10^{275}, as calculated in Appendix A.1, and is not practical to implement. In addition, because PE is based on the GoL, as the number of soups that can be evolved increases, so will the

number of unique patterns created according to the ongoing search for GoL still lifes called Catagolue, where over 7.3 X 10^{15} unique objects have been found, and that number continues to grow.

This interdependence between the number of possible soups and created patterns, in my opinion, recalls the famous remark concerning the debates over uncertainty in the Copenhagen interpretation of Quantum Mechanics: “Shut up and calculate!” (Wikipedia). With that admonition in mind, I created a large dataset composed of 1,000-soup subsets. Each soup in a subset was allowed to evolve for a maximum of 2,000,000 generations. Those soups that did not reach a quiescent end state prior to the maximum evolution limit were discarded from the analysis. The soups that did reach a static end state totaled 820,314, and of this total, prior to examining the static soups for stable patterns, the probability distribution of patterns was unknown and hence uniform[11]. In other words, all 820,314 soups *could* have resulted in the creation of a pattern, but only upon inspection of the static soups was *information* about patterns revealed. That inspection showed that 351,426 soups were nonfunctional, i.e., they did not produce a pattern because the perennial cell remained unbound to canonical cells. However, 468,888 soups were functional and produced stable patterns, a fact only known *after* inspection.

An additional fact was also discovered upon examining the relationship between the number of functional vs. total soups, as shown in Figure 2.

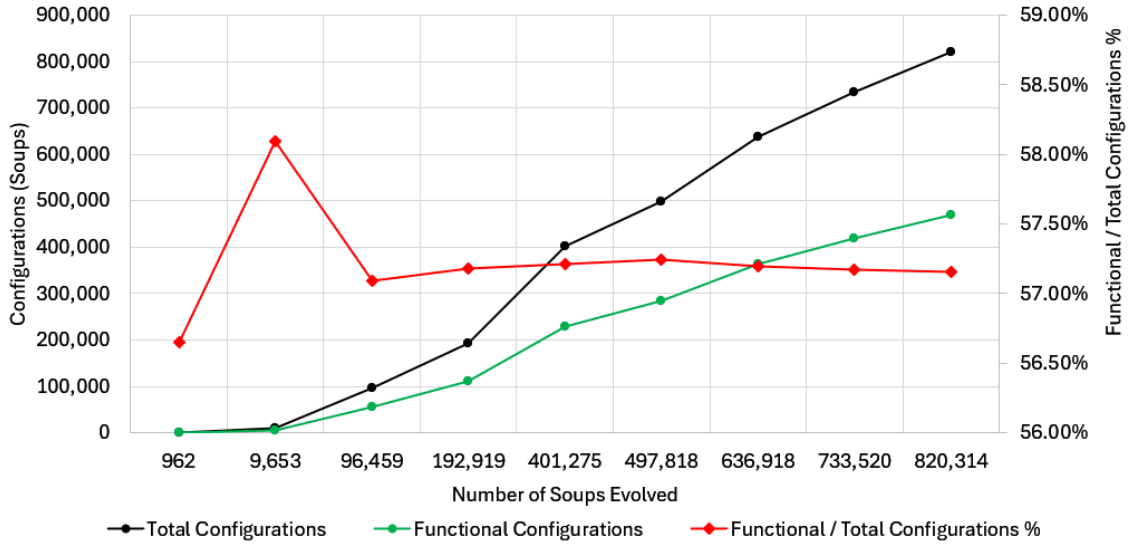


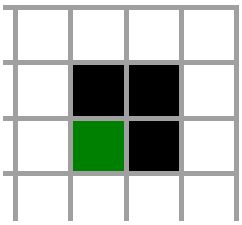
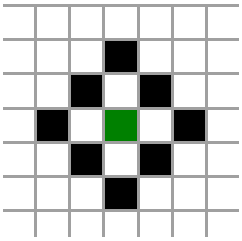
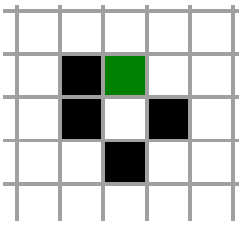
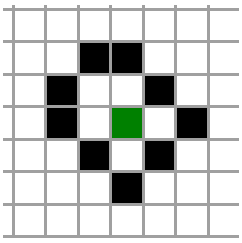
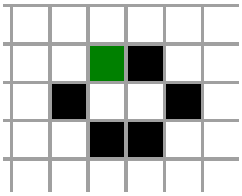
Figure 2. Functional vs Total Configurations. As the number of evolved configurations (soups) grows, the ratio of functional to total configurations is practically constant with a variation of ± 0.05% as the number of soups grows beyond approximately 100,000.

doi.org/10.32388/25T20N

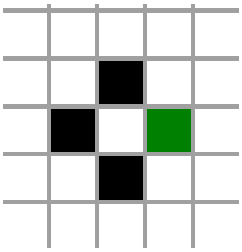
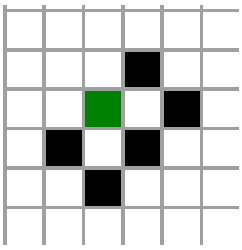
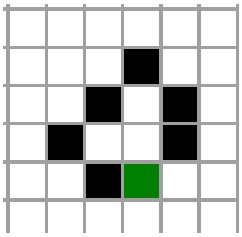
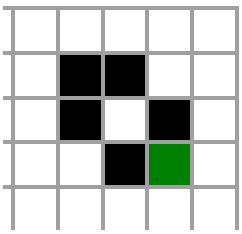
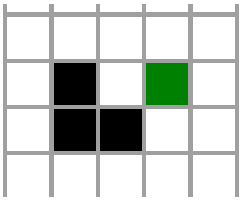
Based on this relatively constant ratio of functional to total soups, the 820,314 soups can be considered a *practical* implementation of all possible configurations. Thus, for any given pattern, the fraction $F(E_X)$ in Equation 1 equals the number of copies of that pattern divided by the configuration size of 820,314. All FI values in bits computed from Equation 1 will thus be *relative* to the configuration size, and the maximum FI value will be set by those patterns with a single copy.

4. Pattern Construction and Functional Information

The dataset produced functional patterns, and consolidation of the number of copies of patterns with more than one form (orientation relative to the CA grid) resulted in 91 perennial patterns with 115,283 copies and 85 canonical patterns with 353,605 copies, for a total of 468,888 copies for both pattern categories. Over 98% of these patterns relative to the total number of copies are shown in Table 1, ranked by their frequency of occurrence in descending order.

Pattern Name	Family	Bonds	Pattern Type	Pattern Geometry	Forms	Copies
Block	3	3	Canonical		1	191,864
Diamond	8	4	Perennial		1	109,411
Boat	4	3	Canonical		4	40,708
Barge Siamese Loaf	9	9	Canonical		4	28,199
Beehive	5	2	Canonical		2	24,087

doi.org/10.32388/25T20N

Pattern Name	Family	Bonds	Pattern Type	Pattern Geometry	Forms	Copies
Tub	3	2	Canonical		1	21,880
Barge	5	3	Canonical		2	17,810
Loaf	6	2	Canonical		4	16,472
Ship	5	2	Canonical		2	7,575
Isomer	3	1	Perennial		8	3,257

doi.org/10.32388/25T20N

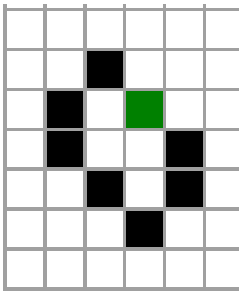
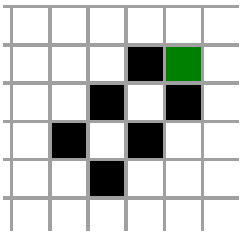
Pattern Name	Family	Bonds	Pattern Type	Pattern Geometry	Forms	Copies
Mango	7	2	Canonical		4	1,083
Long Boat	6	2	Canonical		4	1,066

Table 1. Dominant Patterns. 'Family' represents the number of canonical cells in the pattern and 'Bonds' enumerates the number of canonical cells bound to the perennial cell by the STR. Where a pattern exists in several orientations relative to the grid, the number of 'Forms' of the pattern is listed, and the number of copies represents the total of all forms. The only two perennial patterns are named 'Diamond' and 'Isomer' and the canonical patterns are labeled with their GoL community names.

Several canonical patterns in Table 1 are frequently occurring GoL still lifes but, except for the Block, have a different frequency of occurrence in PE. For example, the Barge Siamese Loaf, abbreviated as BS Loaf in future references, is normally encountered at rank 134 in a GoL soup's evolution but ranks fourth in PE. These variations in rank are due to the different construction dynamics in PE compared with the GoL. All canonical patterns are bound by either 2 or 3 bonds to the perennial cell in keeping with the STR B3/S23 rule, but the two perennial patterns have bonding values of 1 and 4 where the stability of the pattern is maintained by the single perennial cell. Some perennial patterns not shown in Table 1 have bond values of 5 and 6 but are created with a smaller number of copies than the dominant patterns. The family 3 isomer from Table 1 with 8 forms has a geometric structure that occurs frequently in larger families with a greater

number of canonical cells (see examples in Appendix A.2). In Supplemental Data, in Excel spreadsheet format, all the patterns created in PE can be viewed where families of patterns were created up to size 19.

The selection bias demonstrated in Table 1 is a result of the deterministic STR acting on each cell in the PE universe with each of its eight neighbors at each discrete generation of evolution and demonstrates a particular type of contingent behavior. This contingency is described in the context of paleontological history as "...a thing unto itself, not the titration of determinism by randomness"[12] and popularized the idea that rewinding the tape of evolution would create a different history from what is observed in the fossil record. CA have a similar contingent behavior despite being deterministic – change the location of a single cell in a soup and evolution may produce a different result.

In biology, differential reproductive success driven by random genetic mutations is defined as *natural* selection in contrast with PE, where the perennial cell creates an *artificial* selection process due to the catalytic effect of the cell. This process, by attracting canonical cells, forms stable structures in the region of the CA universe in which it is located. This attraction results in changing the dynamics of construction in the area around the perennial cell while the dynamics in other areas of the universe follow the canonical GoL rules. For examples of how the formation of the Block and Diamond – the two most dominant patterns in PE – affects overall evolution due to the perennial cell's persistent active state, see Appendix A.3.

In PE, as analyzed by FI theory, the artificial selection described above is quantified in bits, allowing a relative comparison between patterns as seen in Figure 3 where the FI values of all patterns discovered in the dataset are displayed.

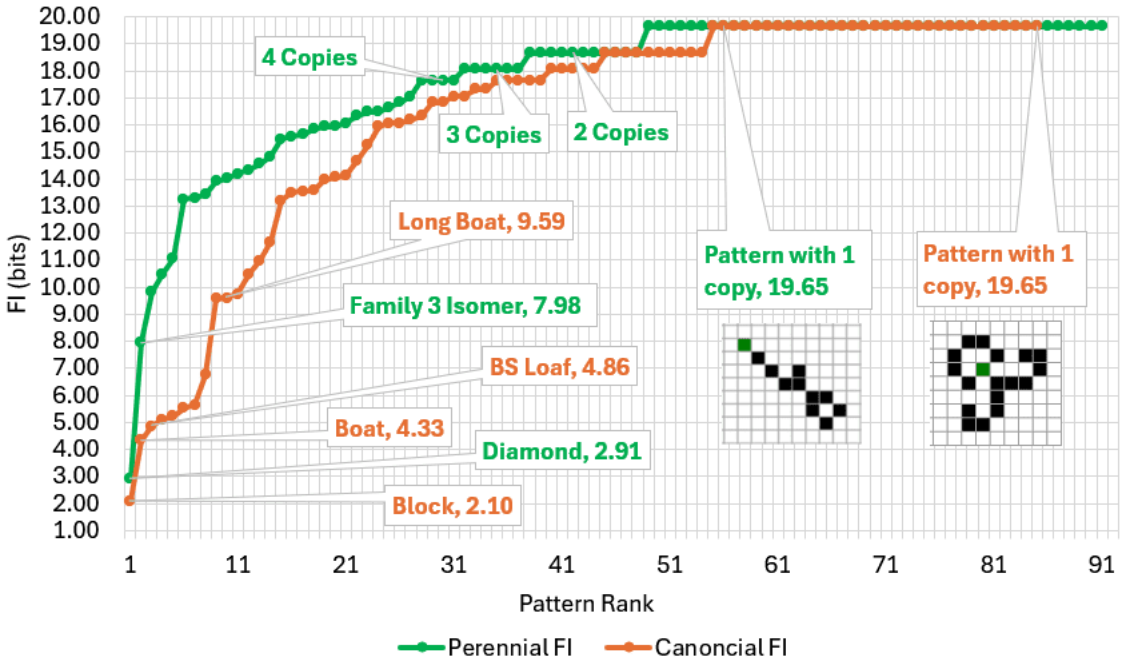


Figure 3. Functional Information of Patterns. Both pattern types reached the same maximum FI value, but the perennial patterns' degree of function rose more quickly relative to rank than the canonical patterns. The "stair-step" feature of each pattern's trend near the larger FI values is due to the small number of copies at these values, as shown on the perennial trend but common to both pattern types. Examples of patterns with a single copy are shown in the lower right portion of the graph. Not all canonical patterns from Table 1 are labeled to reduce clutter on the graph, but both dominant perennial patterns are shown; labels contain the pattern name followed by the FI bit value. Fractional bit values are shown to emphasize the nature of the metric as a continuous measure of relative selectivity among patterns.

The dominant patterns of Table 1, as well as all other patterns, exhibit a rise in their degree of function, and as seen in Figure 3, PE as a complex system displays evidence of increasing functional information over the course of its evolution in compliance with the LIFI.

In PE, if additional soups were to be evolved, new patterns with a single copy would have a slightly higher FI value, but the asymptotic limit would still be reached due to selection for function and the finite size of a new practical ensemble of soups.

4.1. Functional Information and Scale

Since what has been tested so far is based on a grid size of 31 rows and columns, a question of *scale* arises. Are there any differences in the trend of FI values relative to universe size for the created patterns? To

doi.org/10.32388/25T20N

address this question, additional datasets were created and evolved for one smaller and two larger universe sizes with the goal of creating the dominant patterns of Table 1 in the same rank as the primary dataset of 31 rows and columns. The test criteria were created so that the configuration size is limited to the value when each of the universes first produce the dominant patterns in the same rank as the 31 X 31 grid at 820,314 total configurations shown above in Table 1 and Figure 3. This configuration size is termed a threshold value and demonstrates the scalability of FI values as displayed in Figure 4.

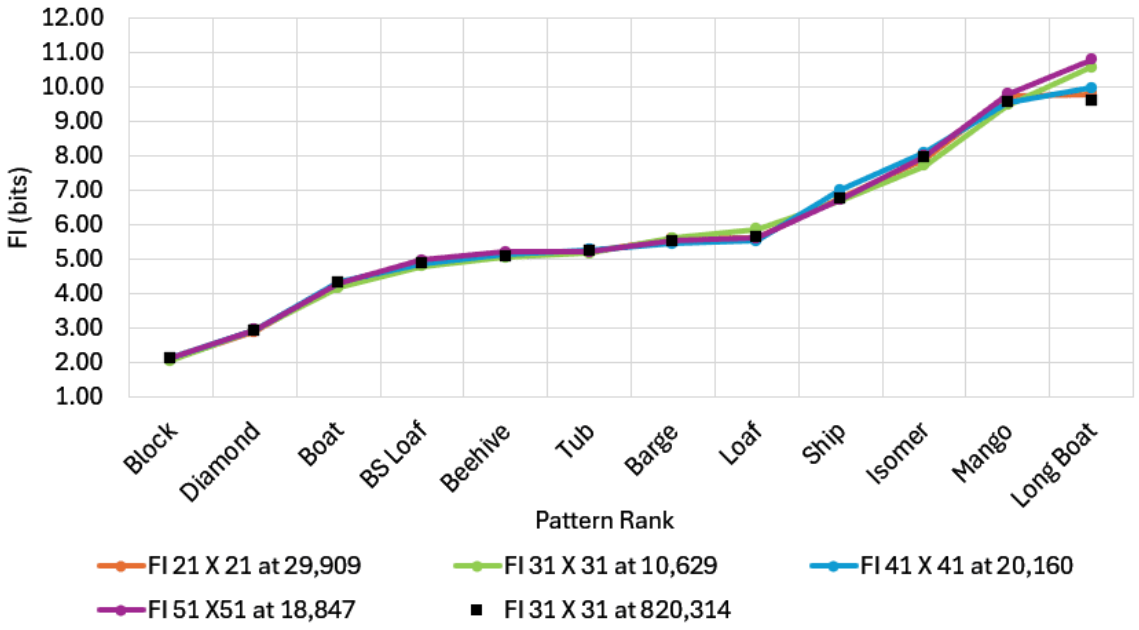


Figure 4. Scalability of Functional Information. The legend for each universe shows the threshold configuration size. The primary dataset of 820,314 configurations shows each pattern as a black square to aid comparison.

Except for the variations between the Mango and the Long Boat, all the universes show nearly identical FI values for the patterns. The slight variation in the threshold FI values for the Mango and Long Boat for all universes is due to the small differences in the number of copies (< 9) for these two patterns due to the threshold limit. For example, in the 21 X 21 universe, the Mango and Long Boat only differ by a single copy, and hence the FI values are nearly identical. In the plot at 820,314 configurations, the Mango has 17 more copies than the Long Boat; thus, the difference in rank between these two patterns is small despite the large size of the primary dataset's configuration. This fact does not invalidate the scalability of the universe size: the selective construction of dominant patterns occurs on all universe sizes as a function of the STR operating identically without regard for universe size, with one exception: there is a lower limit to the universe size that would restrict the formation of some patterns. This restriction would be a result of

not having enough space in a universe for patterns that exceed the size of the grid. For example, in the primary database of 31 X 31, a few patterns were as large as 20 cells, and thus a grid smaller than this value would not allow such a pattern to form.

There is another aspect of scale that might at first glance appear incongruous: the size of the threshold configurations varies greatly among the four universes in Figure 4. If the four tested universes were each created with the same number of configurations, a difference in FI values would be more noticeable, as shown in Figure 5.

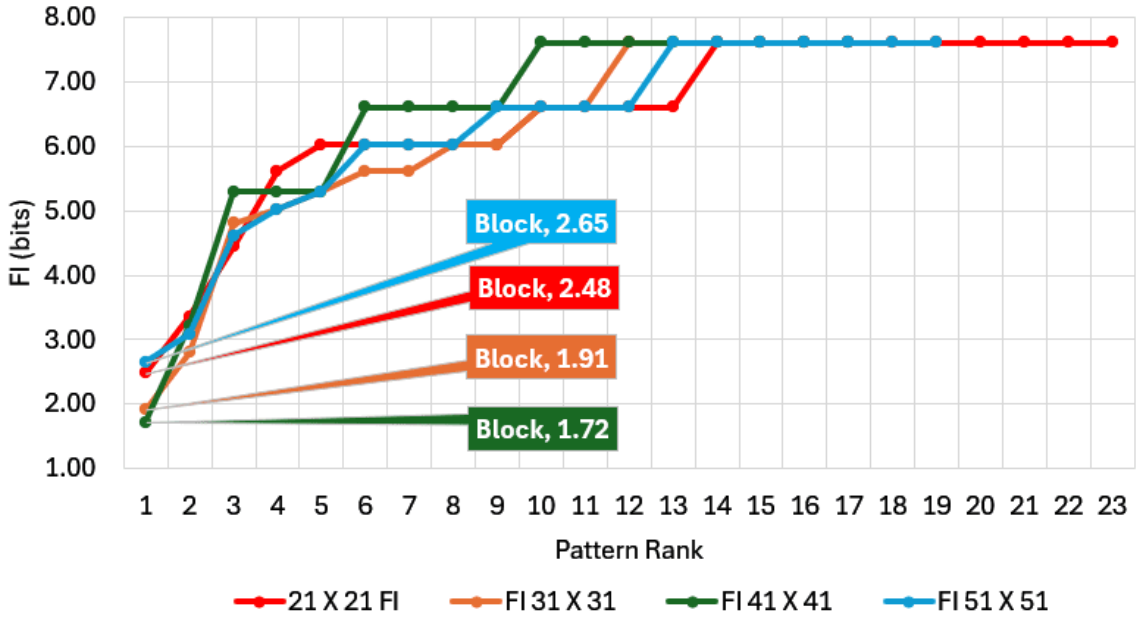


Figure 5. Evolution with 195 Configurations per Universe Size. This was the first increment of configurations toward achieving the threshold limit as shown in the previous figure. Not all the dominant patterns were created with this configuration size.

What is demonstrated by Figure 5 is that universe size does affect the number of configurations required to achieve a given rank of pattern creation. This is due to the nature of random soups where, in addition to the dominant patterns being created as shown in Figure 4, many other patterns were created, and more soups had to be evolved to achieve the same rank as the primary dataset. For example, despite all four universes in Figure 5 having the same configuration size, the number of patterns created differed from 24 being created at 21 X 21, 19 at 31 X 31, 18 at 41 X 41, and 20 at 51 X 51.

These facts provide information relevant to model design, since universe size also affects the computational workload of evolving soups to their quiescent end state. For example, the software created

doi.org/10.32388/25T20N

for this project processed each soup in a configuration at a rate of 10 seconds/soup for the 21 X 21 universe and 25 seconds/soup for the 51 X 51 universe, while the primary dataset in the 31 X 31 universe at the configuration size of 820,314 soups took 15 seconds/soup and expended 3,477 total hours running on a parallel cluster of high-end desktop computers. When computation time is a factor, either due to software design, hardware capability, or availability, these observations inform decisions about model design with confidence in the chosen size's results scaling when a different choice is made.

4.2. Functional Information, Construction, and Computational Work

As seen in Figure 3, the Block at an FI of 2.10 bits requires less information, compared with other patterns, to achieve a functional state. In this case, can less information be equated with the amount of computational work performed by PE to create the Block than the Diamond at 2.91 bits? If a correlation between information and computational work exists, this suggests a hypothesis that fewer generations of evolution of a soup are required to construct a Block compared with other patterns.

To test this hypothesis, the software created for this project was designed to record the minimum, maximum, and median number of generations for each pattern's creation for each soup in which it was generated, and is defined as construction steps (CS). The CS values are different from the number of generations that a random soup requires to reach its end state; rather, it is a measure of the generational value when the pattern first forms and remains a strict still life until the end state is reached, and is a measure of the computation work done by the STR on a random soup. Figure 6 shows an example of these CS values for the Block.

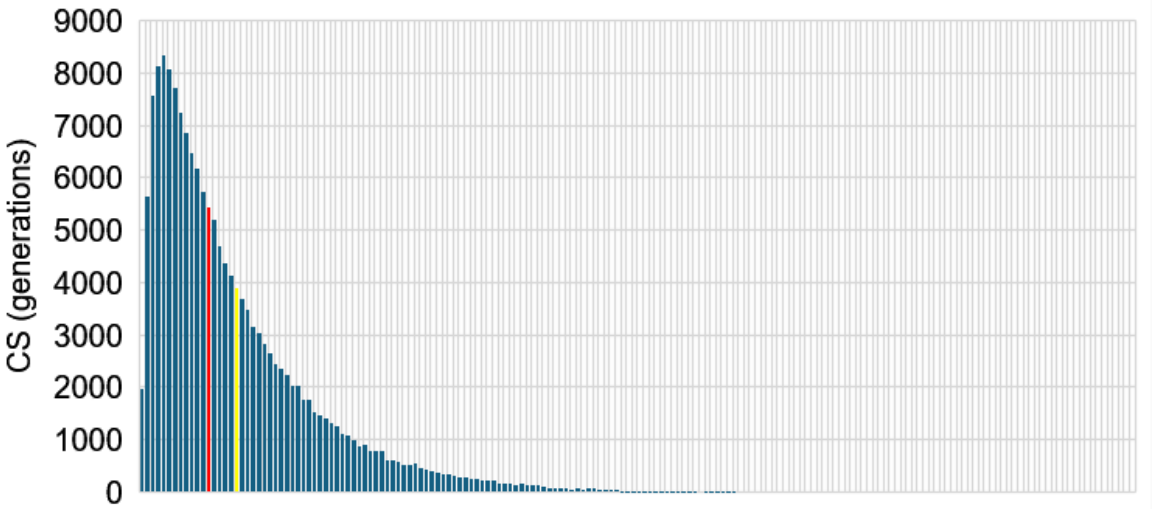


Figure 6. The Block's Construction Steps. For all 191,864 copies, the median value of 546 generations is shown as a red bar, and the average at 748 in yellow. The minimum CS value was 4 generations and the maximum was 7,649. The 178 bin values in this histogram are not shown due to space limitations on the page.

The range of CS values has a long tail, common for CA evolution where a power law is active during the self-organizing phase[13]. The shape demonstrates the occurrence of "black swan events"[14] in the long tail, that while less numerous, distort the validity of the average as a *typical* value. For this reason, the median value is more representative of a typical value[15] but only becomes reliable as a typical value when many copies exist for a pattern as seen in Figure 7 where several of the Block's incremental median CS values are displayed.

doi.org/10.32388/25T20N

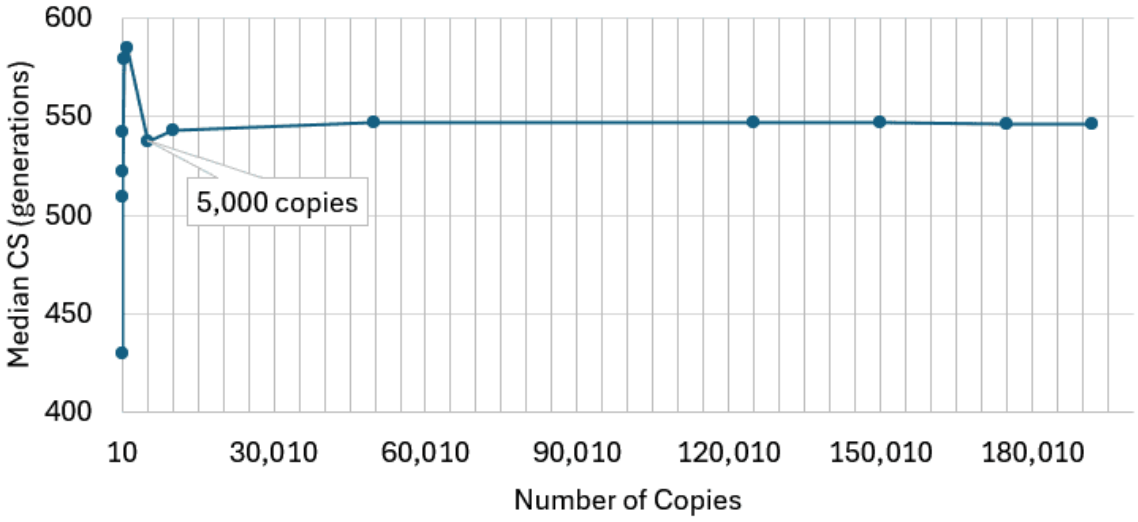


Figure 7. Incremental Measurements of the Block's Median Construction Steps. After 5,000 copies, the median value begins to converge to its final value of 546 generations at 191,864 copies.

In Figure 8 the median CS values of the 12 dominant patterns of Table 1 are displayed along with patterns ranked from 13 to 20 that had a smaller number of copies.

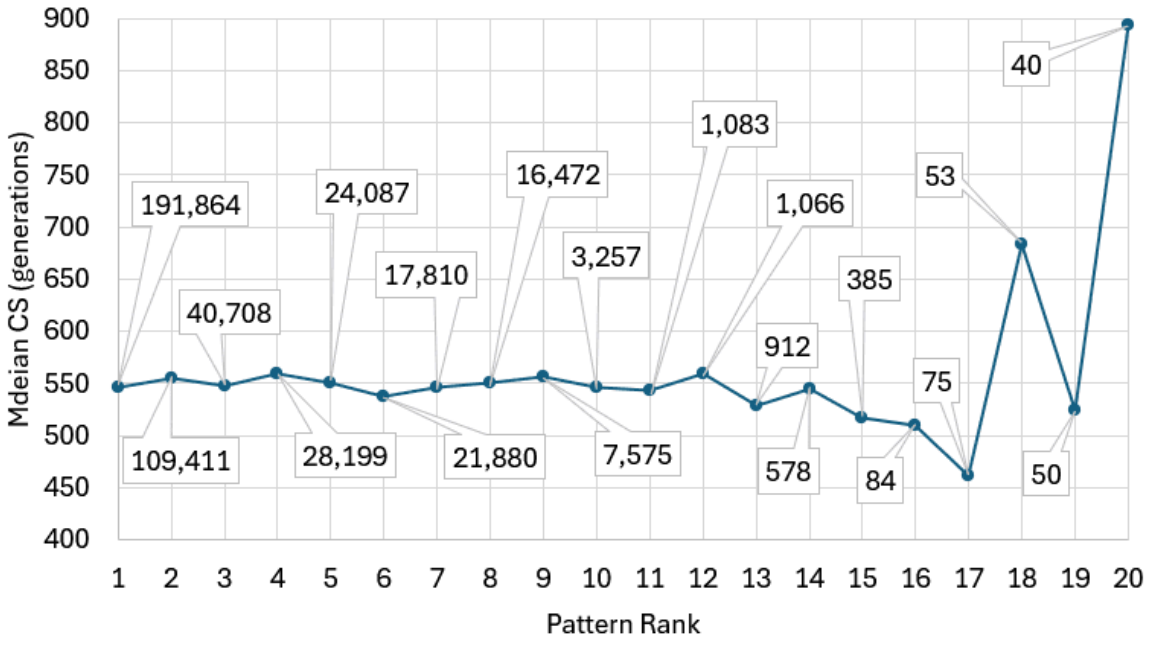


Figure 8. Median CS of 20 Patterns. The datapoints are labeled with each pattern's number of copies.

doi.org/10.32388/25T20N

The average median value for the dominant patterns of Table 1 (rank 1 to 12) is 550 ± 6 generations, a remarkably small variation considering the differences in FI values for the patterns. From rank 13 to 20 the number of copies ranges from 912 to 40 copies and have a median variation of ± 87 generations.

When the median CS is normalized as a ratio by a pattern's number of copies it defines a new metric called the Construction Index (CI) representing the typical number of construction generations per pattern copy. This normalization method is justified since a *single soup* is the origin of a *single copy* for a specific pattern.

As an example, the Block's normalized CS value is calculated as follows:

CI_{Block} = Median CS_{Block} / $Block_{copies}$ = 546 / 191,864 = 0.0028 generations/copy

For the dominant patterns of Table 1, when the CI values are compared to the FI values, a correlation in trends is observed and displayed in Figure 9.

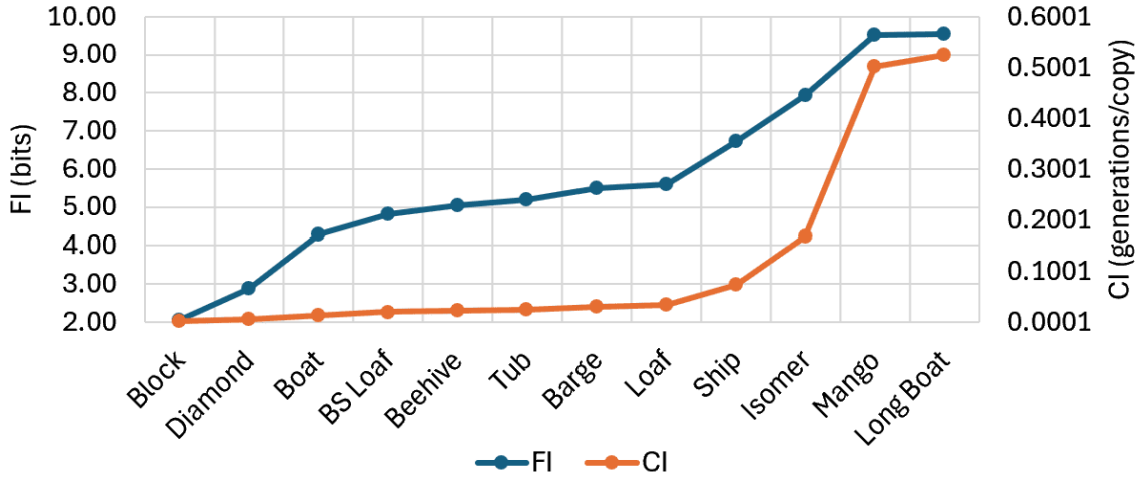


Figure 9. FI and CI Trends. Since the FI and CI values are incommensurate, their values must be represented on separate axes, but the trends of the two metrics are similar. As more information is required to construct a pattern, a larger number of generations is expended to create the pattern from a random soup.

The hypothesis proposed that information and construction are linked has validity within limits governed by the number of copies of a pattern. As the number of copies of patterns decreases the median value variations increase, resulting in less correlation between the FI/CI trends as seen in Figure 10.

doi.org/10.32388/25T20N

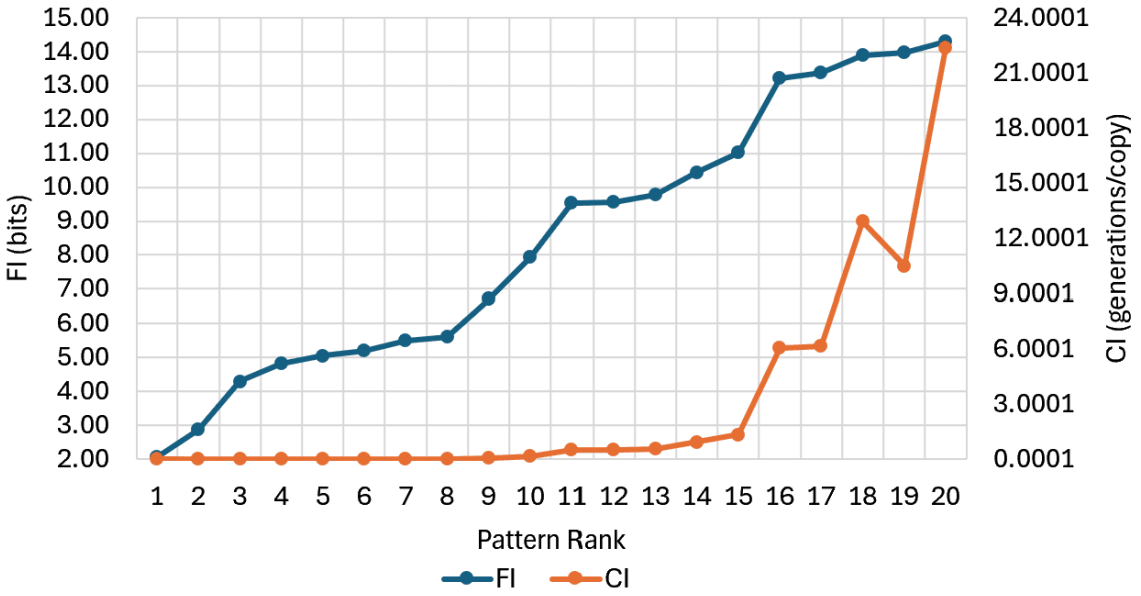


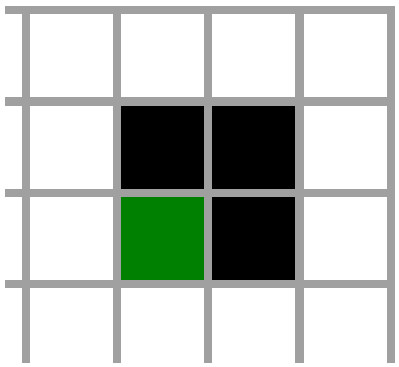
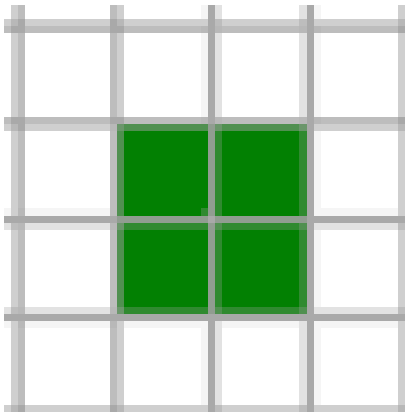
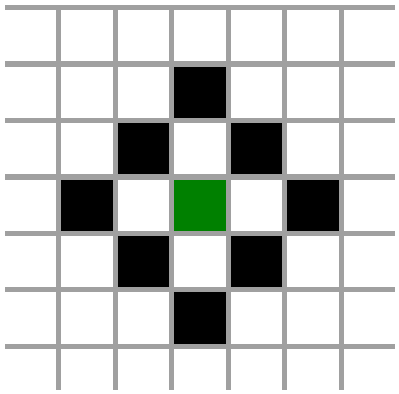
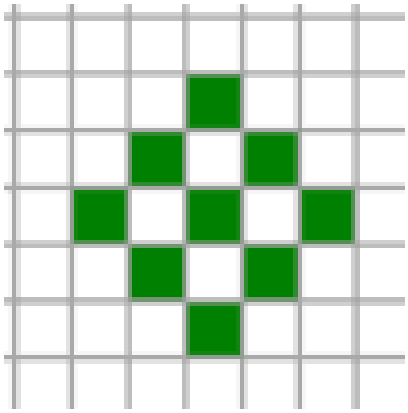
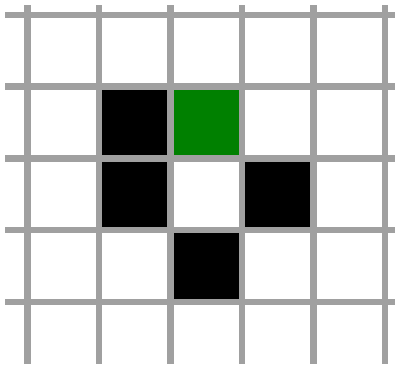
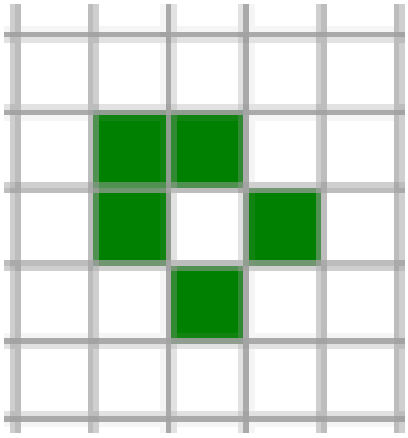
Figure 10. Degradation of FI/CI Correlation. The standard deviation of the median values from rank 13 to 20 is ± 87 generations as compared with ranks 1 to 12 with a variation of only ± 6 generations.

The reader may have questioned if the term Construction Index is related to the Assembly Index (AI) of the theory invented by astrobiologists to aid in detecting possible biosignatures in the spectroscopic signals from the atmospheres of exoplanets. A comment is warranted since this author has collaborated in NASA's astrobiological research[16][17] and the CI terminology was inspired by that work. Assembly Theory, with its origin in astrobiology[18], has been recently modified to apply to cellular automata[19] and is called AssemblyCA. In the application of this theory to CA, objects – called patterns in my current work – are built from preexisting objects and the assembly of these objects into more complex objects is directed by an agent. The AssemblyCA of this theory is a measure of the shortest number of steps to construct a more complex object from simple ones. The theory has merits within the context of its experimental design but is different from the design of the PE model that has only random configurations creating patterns without any supervisory agent. To the interested reader, I recommend a look at this work for another method of complex systems analysis.

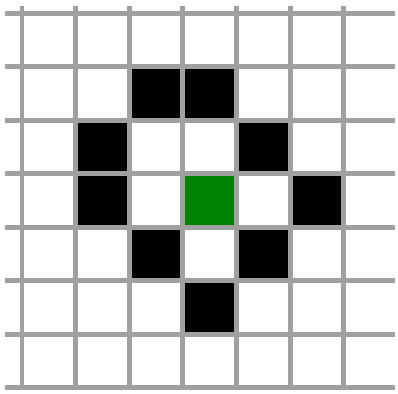
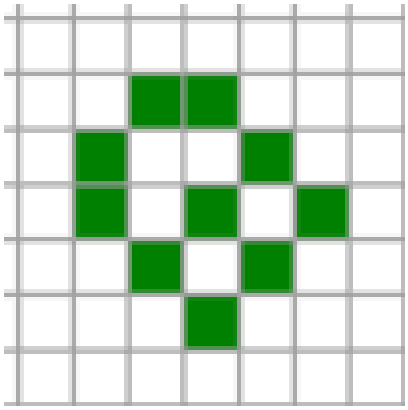
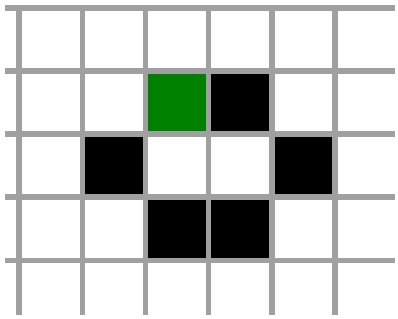
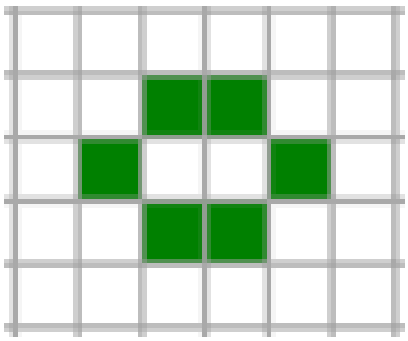
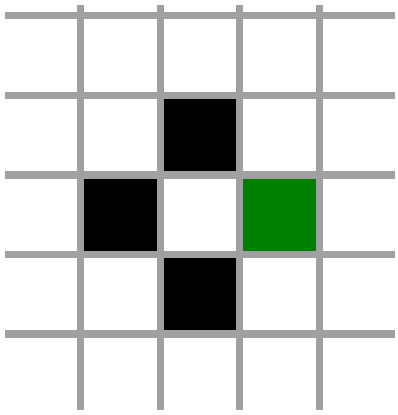
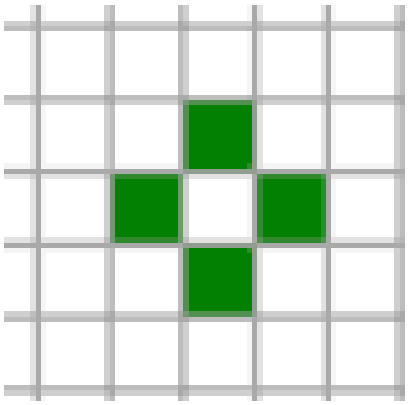
5. Phase Transitions and the Growth of Complexity

The patterns presented in the previous section can be considered as analogous to how the primordial elements created by stellar evolution in the early universe[20] eventually gave birth, over cosmological ages,

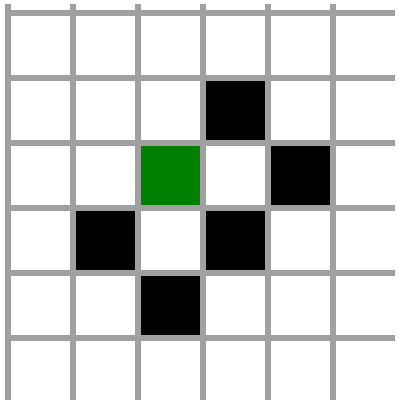
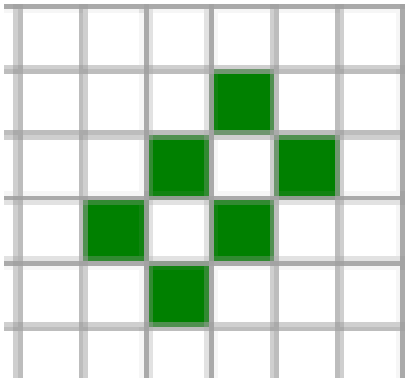
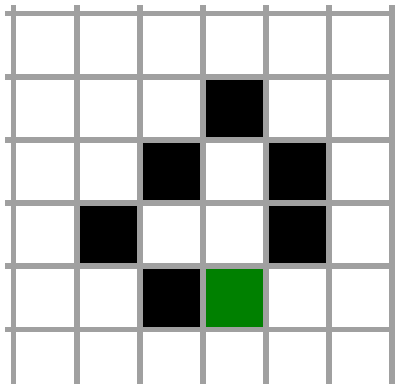
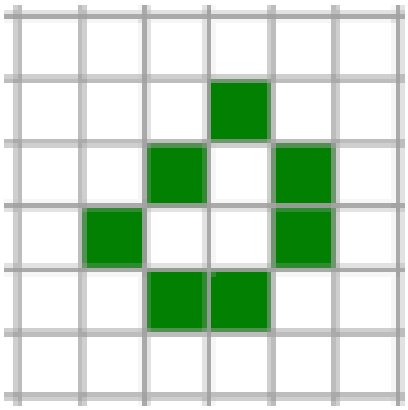
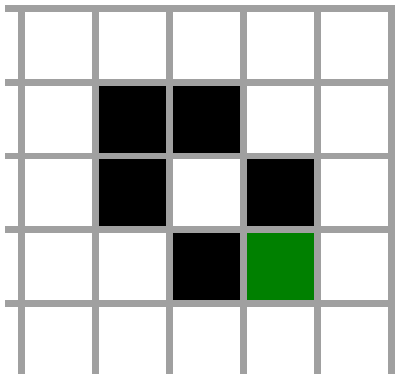
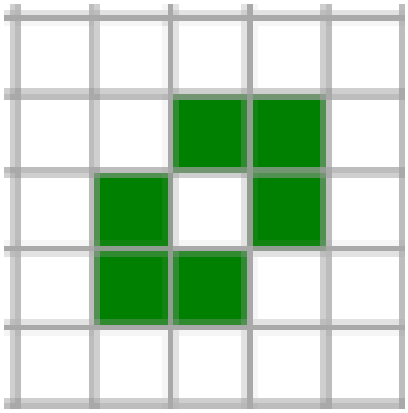
to minerals that formed the geospheres of planets[21]. These patterns demonstrated increasing FI over the course of their evolution and will be referred to as phase 0 patterns. In this section the phase 0 patterns transition to a new form analogous to minerals that have been subjected to thermodynamic and gravimetric churning of the geosphere[22] in continuing phases of structure building[23]. This new form will be modeled by converting all canonical cells of a pattern to perennial cells as shown in Figure 11 for the dominant patterns from Table 1.

Pattern Name	Phase 0 Geometry	Phase 1 Geometry
Block		
Diamond		
Boat		

doi.org/10.32388/25T20N

Pattern Name	Phase 0 Geometry	Phase 1 Geometry
BS Loaf		
Beehive		
Tub		

doi.org/10.32388/25T20N

Pattern Name	Phase 0 Geometry	Phase 1 Geometry
Barge		
Loaf		
Ship		

doi.org/10.32388/25T20N

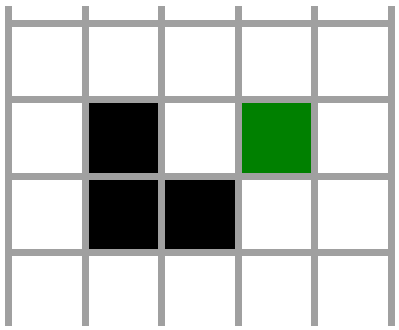
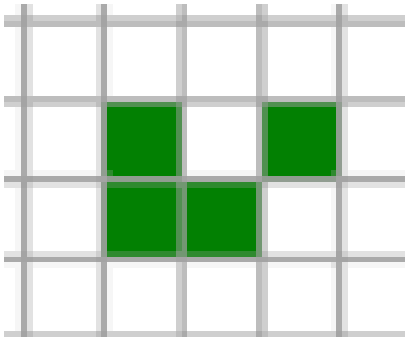
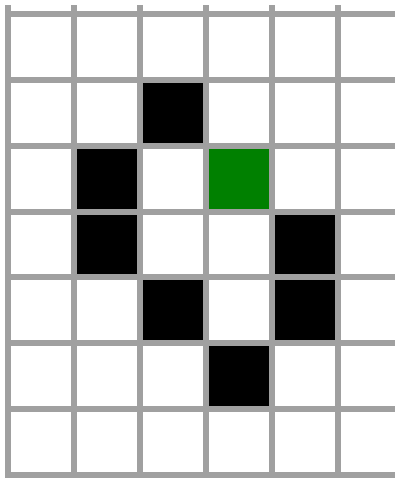
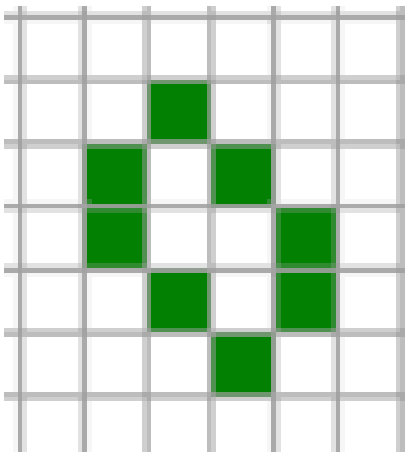
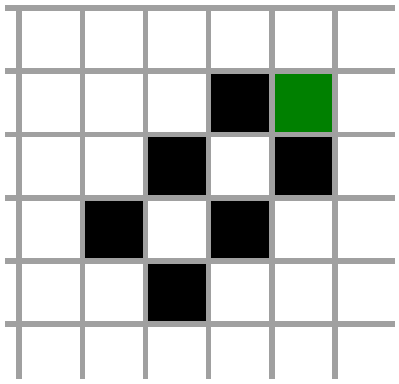
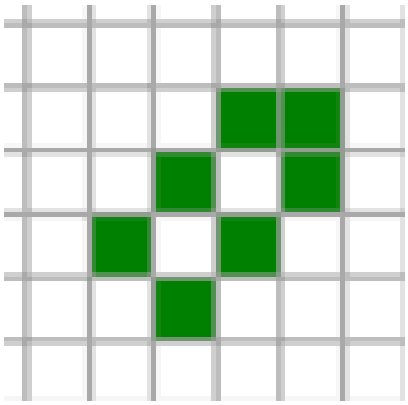
Pattern Name	Phase 0 Geometry	Phase 1 Geometry
Isomer		
Mango		
Long Boat		

Figure 11. Phase Transition of Patterns from Phase 0 to Phase 1.

doi.org/10.32388/25T20N

Phase 1 patterns function as seeds for potential growth into new forms by interaction with random soups of canonical cells, just as phase 0's single perennial cell interacted to create these seeds. Each perennial seed exists as a new system of PE with dynamics driven by its size and geometry. When compared to phase 0's single perennial cell, each of the seeds in Figure 11 injects more order into an evolving soup, extending the nonequilibrium phase of structure building. This has the effect of increasing the computational workload of producing patterns. For example, the primary dataset's number of cycles for phase 0 took an average of 900 cycles/soup to process a soup, while the phase 1 seeds averaged 18,000 cycles/soup—see Appendix A.4 for details.

Evolutionary results for phase 1 seeds are shown in Figure 12, where each seed displays an increase in functional information as predicted by the LIFI, but there is large variability in the FI increase per seed as well as in the number of patterns created.

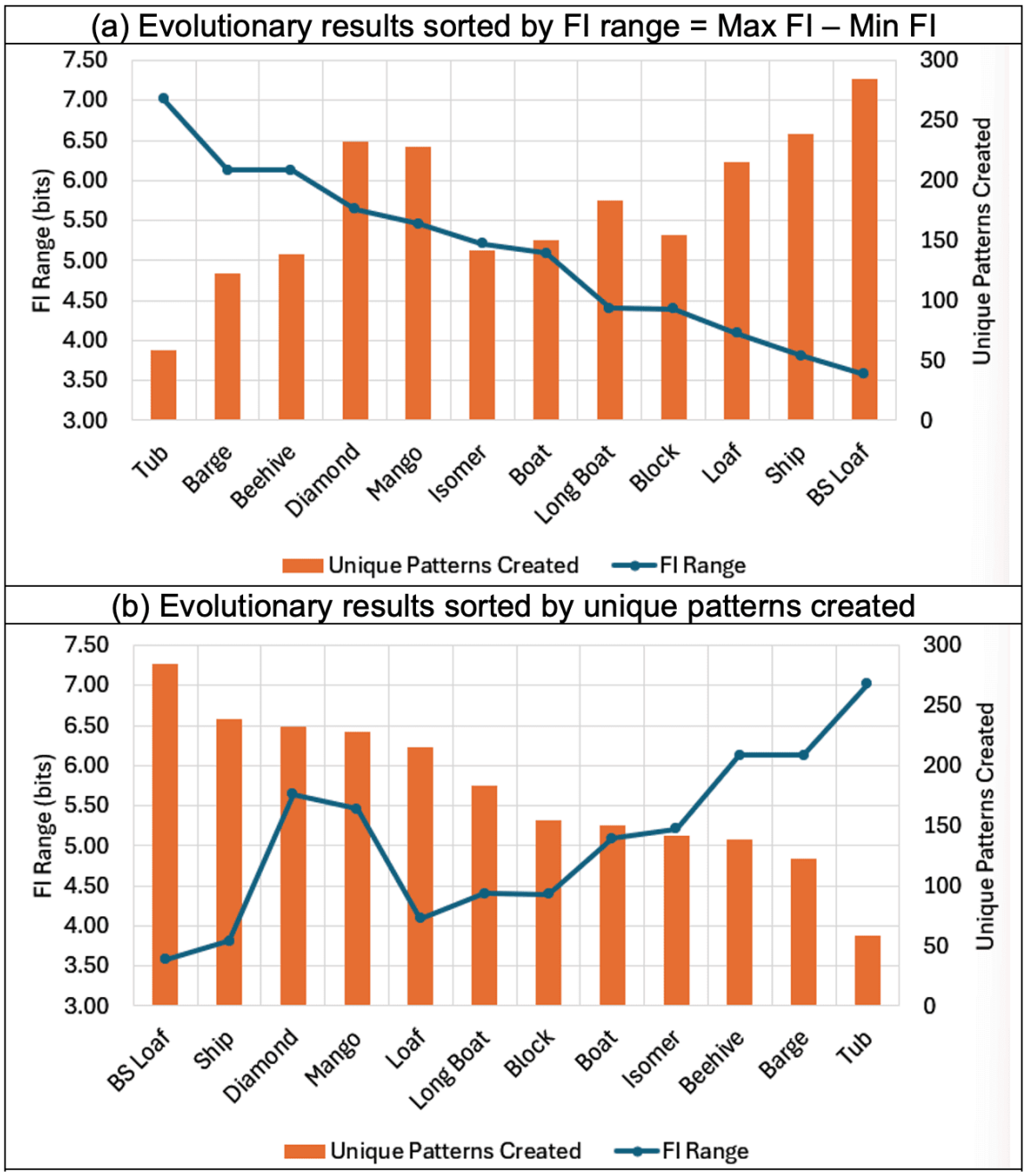


Figure 12. Two Views of Phase 1 Evolutionary Results. In (a) and (b), the results are sorted in descending order by the two metrics, respectively. The average configuration size was 536 ± 17 soups for each seed. By comparison, phase 0's single perennial cell had an FI range of 7.14 bits and created 28 unique patterns when evolved for 536 soups.

The two views of Figure 12 allow a partial correlation to be made between the FI range and the number of created patterns. The BS Loaf and Ship have the smallest FI range but created the largest number of

doi.org/10.32388/25T20N

patterns, while the Tub, Barge, and Beehive with large FI ranges created the smallest number of patterns. In this case, the Tub, with the largest FI range, created the smallest number of patterns, followed by the Barge and Beehive, whose FI range was almost identical. This is the limit of any direct correlation between FI range and pattern creation, but a connection does exist and is related to the selectivity of patterns created by a seed, as seen in a view of the two extremes, i.e., the Tub vs. the BS Loaf, demonstrating that a greater degree of selectivity is exhibited by the Tub than the BS Loaf, as shown in Figure 13.

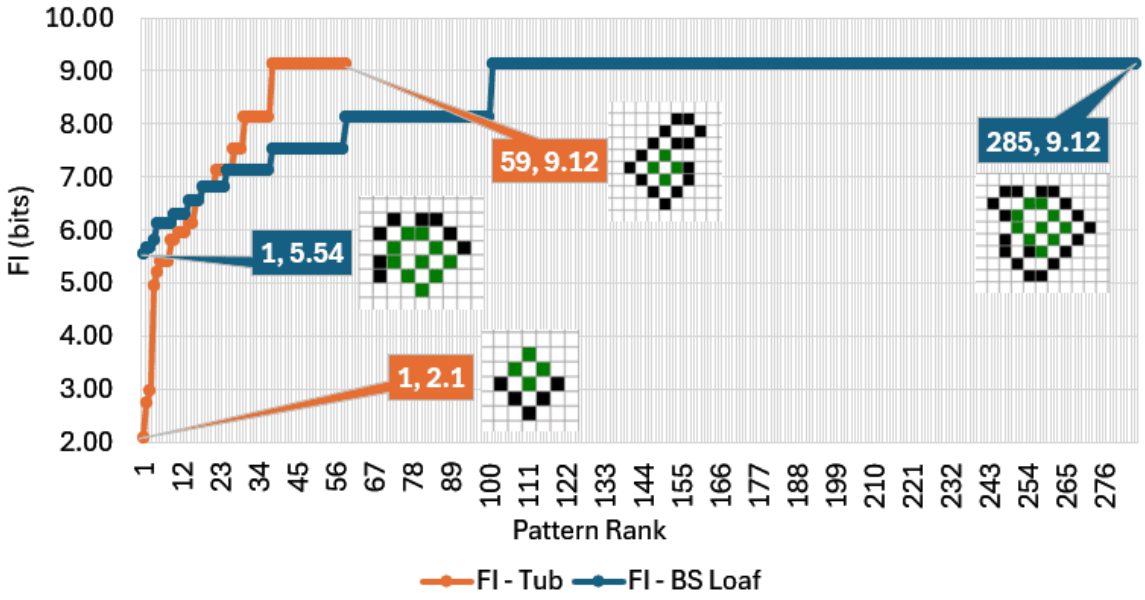
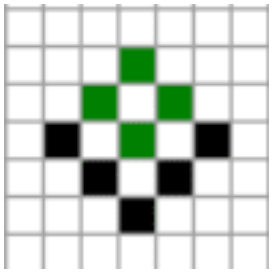
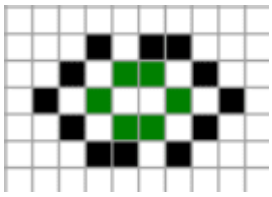
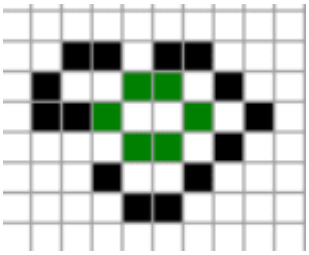
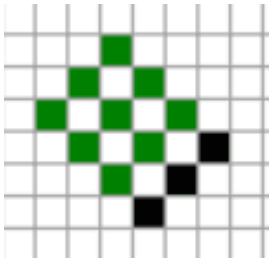
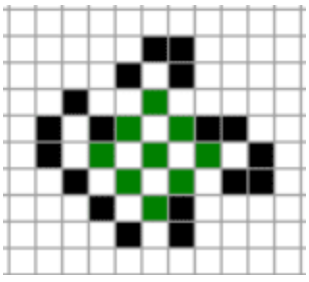


Figure 13. FI Range vs. Number of Created Patterns for the Tub and BS Loaf. The configuration size of these two systems with different seeds is identical at 557 soups. The Tub generated 59 unique patterns, and the BS Loaf generated 285. The inset images are the patterns at the minimum and maximum FI values, and the label for each shows the pattern's rank followed by its FI value.

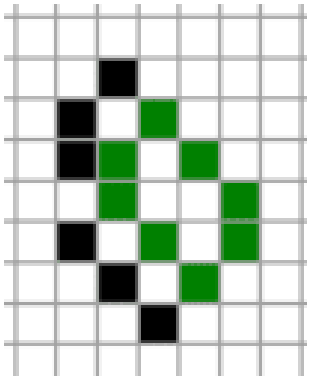
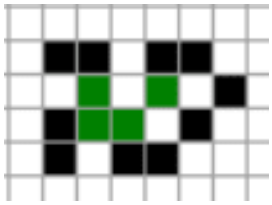
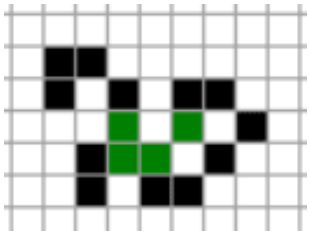
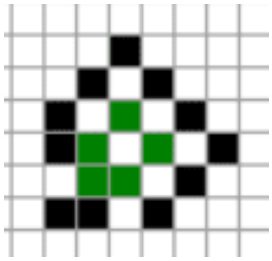
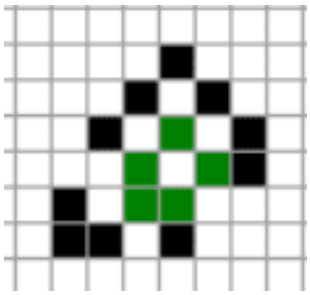
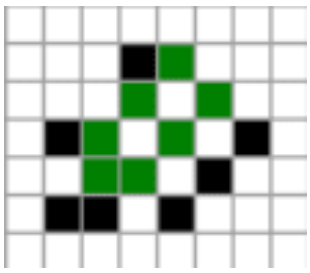
The growth of complexity as defined by FI theory is apparent for both seeds, but the greater selectivity of the Tub is due to the creation of a larger number of patterns at a lower FI value—indicating a larger number of copies per pattern—than the BS Loaf at its lower FI values. Thus, "greater selectivity" means the Tub spent more evolutionary time creating copies of low-FI-value patterns than patterns with a higher value.

In Table 2, examples of created patterns at the minimum and maximum FI values for each seed in phase 1 are presented. The table is sorted by the FI range in descending order, and the number of unique patterns created during the evolution of each seed is listed.

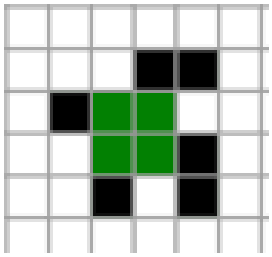
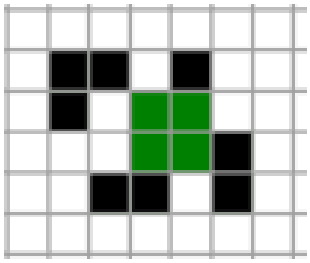
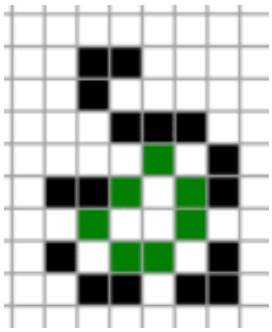
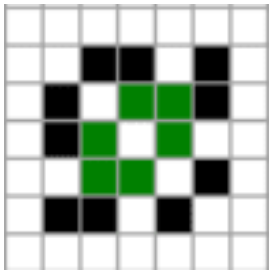
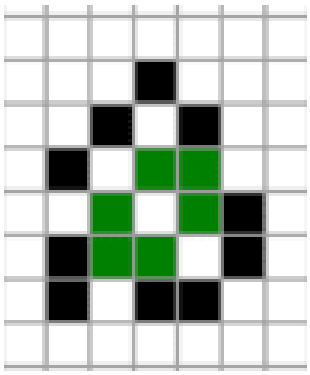
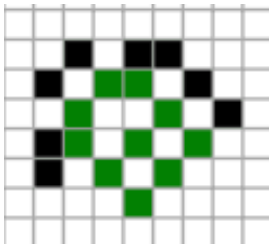
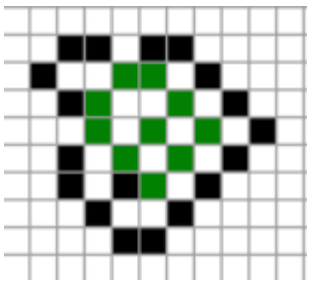
doi.org/10.32388/25T20N

Seed	Min FI	Min Pattern	Max FI	Max Pattern	Total Patterns Created
Tub	2.10		9.10		59
Barge	2.94		9.07		122
Beehive	2.89		9.02		138
Diamond	3.36		9.00		232

doi.org/10.32388/25T20N

Seed	Min FI	Min Pattern	Max FI	Max Pattern	Total Patterns Created
Mango	3.64		9.10		228
Isomer	3.81		9.02		142
Boat	3.91		9.00		150
Long Boat	4.66		9.06		183

doi.org/10.32388/25T20N

Seed	Min FI	Min Pattern	Max FI	Max Pattern	Total Patterns Created
Block	4.65		9.04		155
Loaf	5.02		9.11		215
Ship	5.31		9.12		239
BS Loaf	5.54		9.12		285

doi.org/10.32388/25T20N

Table 2. Phase 1 Pattern Examples. The pattern created at the minimum FI value is ranked first in frequency of occurrence. For the Tub and the Barge, the #1 ranked pattern is a Diamond, created in phase 0 as the #2 ranked pattern from a single perennial cell.

The observation that a pattern with a Diamond geometry was created in phase 0 at rank 2 (Table 1) and in phase 1 by the Tub and Barge at rank 1 (Table 2) demonstrates the construction bias of PE toward structures with a diagonal orientation to the grid. This fact, coupled with an actual fully perennial Diamond seed in phase 1 creating at rank 1 another similar diagonal structure, confirms this construction bias, and for phase 2 evolution, this bias continues, as shown in Figure 14.

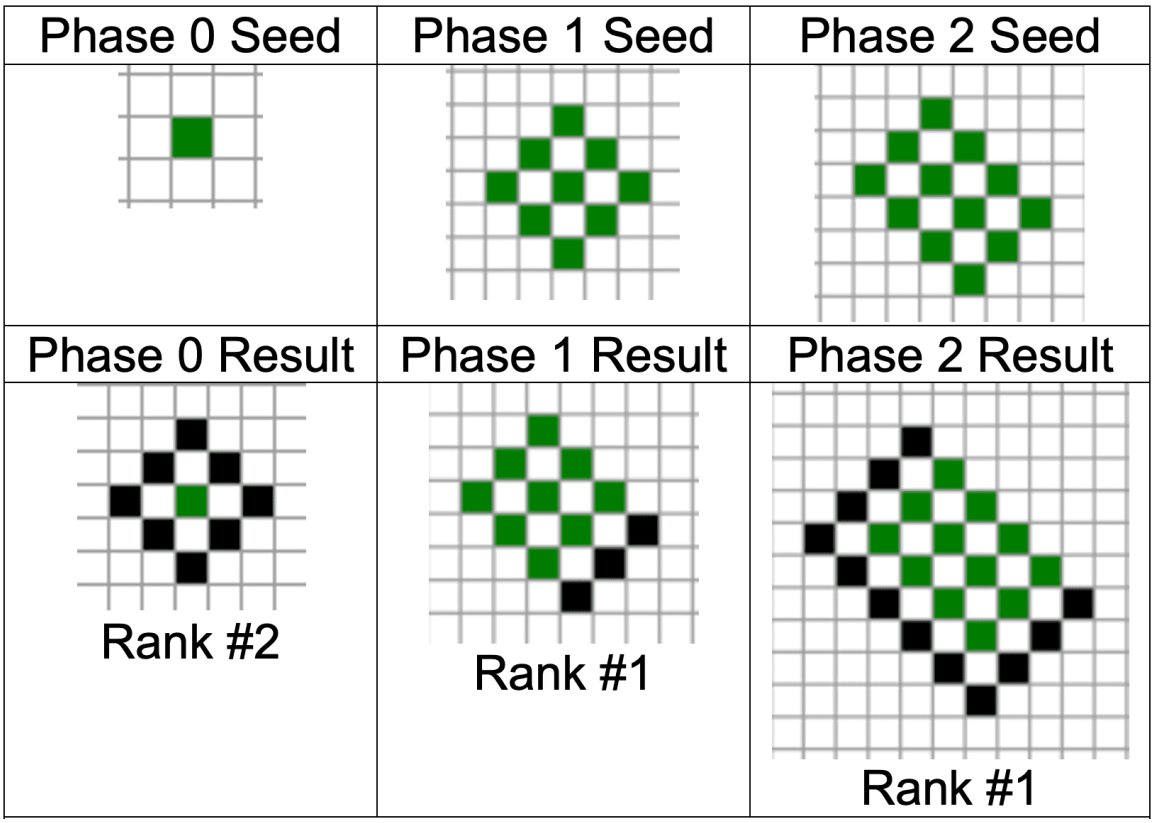


Figure 14. Selection Bias in PE. Binding cells in a diagonal fashion is common. The pattern resulting from phase 1 is made fully perennial as a seed for phase 2.

The rise in FI of a Diamond from its origin in phase 0 to its subsequent structures in phase 1 and phase 2 is shown in Figure 15.

doi.org/10.32388/25T20N

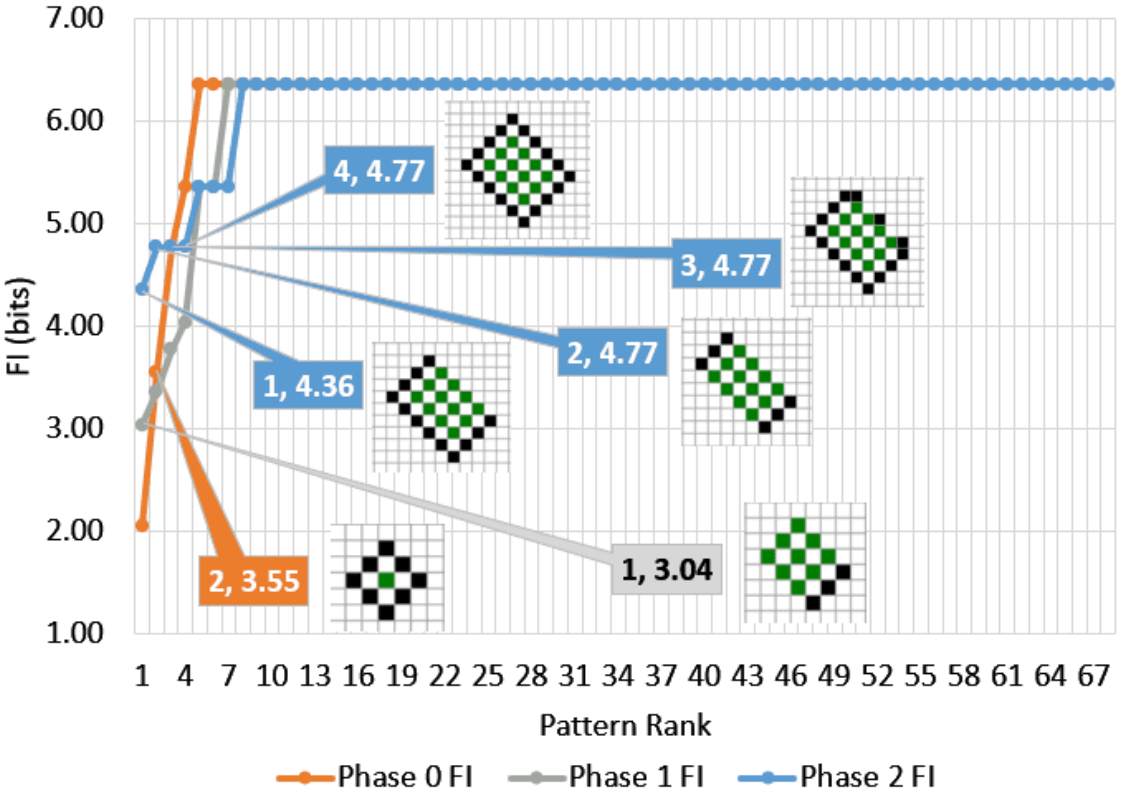


Figure 15. Three Phases of Evolution. The labels contain the pattern rank followed by the FI value in bits. Each phase has a configuration size of 82 random soups. The number of created patterns ranges from 13 at phase 0, to 51 at phase 1, and to 68 at phase 2. Since all phases were run with the same size configuration, the maximum FI value for each phase is the same for those patterns with a single copy.

In phase 2, an increased injection of order due to the larger seed size results in an average of 145,963 cycles/soup to evolve a soup to discover stable structures, and only 82 soups out of an initial 640 resulted in a stable, functional pattern. The construction of diagonal-oriented patterns continues in phase 2, as shown by patterns at rank 2, 3, and 4 with the same FI value.

6. Discussion and Outlook

The LIFI described in this article has been validated for the driven CA model defined in this research in several phases of its evolution. The model, as a minimalistic complex system with a deterministic outcome for each configuration of the system, allowed patterns to be analyzed for their function in maintaining stability and quantified that function as information defining the relative complexity of patterns. Rather than judging the relative complexity of two patterns according to visual impressions of their geometry, FI

theory yields values in bits, providing a numerical ranking of complexity. An observer can now say that one pattern is twice as complex as another without it being a matter of opinion in the eye of the beholder. But beyond quantifying complexity, the correlation between the computational work of construction and the resulting frequency of occurrence of a pattern – when a sufficiently large sample size exists – was a new discovery about cellular automata.

The challenge of PE as a model system, one that demonstrates open-ended evolution, was addressed by the statistical behavior of a large collection of random soups where those soups that produced functional patterns maintained a consistent ratio to the nonfunctional soups. This discovery was important in the context of FI theory, where all possible system configurations must be considered, and allowed a dataset of a practical size to be evolved and produce relative measures of functional information that scale over different universe sizes.

PE, as an artificial life model, demonstrated a selection bias toward preferred patterns with a wide diversity of geometrical configurations. While FI theory can classify these patterns in terms of frequency of occurrence, the STR governs how construction produces the bias. The connection between construction bias and functional information could be explored in more detail with larger configuration sizes. Also, the observation that perennial patterns rose more quickly in complexity in phase 0 compared to the canonical patterns, as measured by their corresponding FI values, could be analyzed in future research to discover more insights into the reasons driving this behavior.

Another source of complex systems with features like PE is biological systems identified as having dynamic kinetic stability (DKS) attributed to the origin of function and complexity in nature[24]. This same scientist has also popularized his research in a book for a general audience entitled “Life’s Chemical Secret: The Origin of Purpose and Mind”[25]. The concept of DKS describes the nonequilibrium state maintained by homeostasis in living cells that shares some characteristics with the extended dissipation region of driven CA where structures are built, with the exception that no self-replication of artificial life occurs. In preliminary investigations in PE, phase 1 patterns, when placed near each other in an empty universe, can result in creating canonical cells interacting over many generations, eventually binding the patterns together into a new structure. This is one example of how this model could be extended in future research.

Appendix

A.1. All Possible Soups

The number of unique soups is the number of ways to choose 336 active cells from 961 cells, given by the formula:

$$\text{Unique soups} = 961!/336!(961-336)!$$

Reference: https://en.wikipedia.org/wiki/Combination

Using Stirling's approximation for large factorials to solve the "Unique soups" above:

$$\text{Unique configurations} = n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

Reference: https://en.wikipedia.org/wiki/Stirling%27s_approximation

$$\text{Unique configurations} = 5X10^{275}$$

A.2. Isomers in Families

The family 3 isomer shown in Table 1 has 8 forms, and other families contain a similar geometric construction as seen in Figure A2.1.

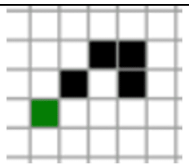
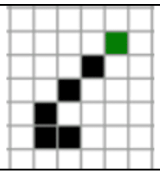
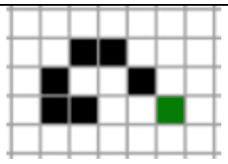
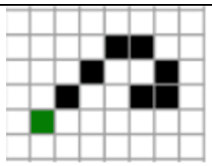
Family	4	5	6	7
Structure				
Forms	8	8	6	2
Copies	578	50	53	2

Figure A2.1. Other Isomers. The number of forms indicated represents those that were detected; other forms might exist for those with 6 & 2 forms with the evolution of additional soups. All of these shown exist with fewer than 600 copies as compared with the dominant isomer from family 3 with 3,257 copies.

A.3. Dynamics of Construction

An example of the attractive effect of a perennial cell is demonstrated during the construction of the canonical Block and perennial Diamond, the two most frequently occurring patterns. Using a heat map designed for this project's software to track the state transition (active-to-inactive or inactive-to-active) of each cell in the CA universe at each generation of evolution, a measure of the number of generations of evolutionary time a cell is active illustrates this attraction. In both heat maps, the soup with and without the single perennial cell evolved to an end state, displaying the contingency inherent in CA, as the change in a single cell's initial state makes a significant difference to the end state of the soup.

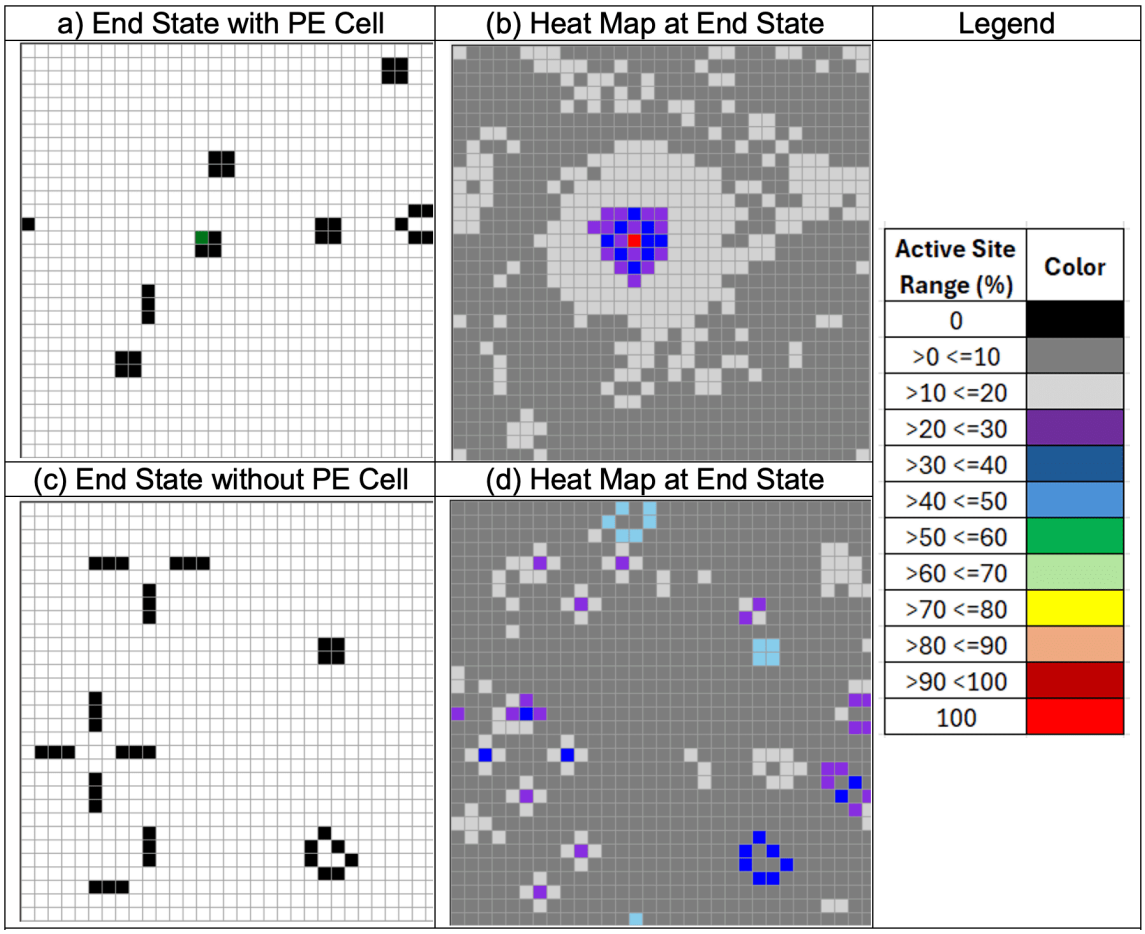


Figure A3.1. Heat Maps for the Block. The end states of evolution with and without the PE cell are shown respectively in (a) and (c). In (a), the Block was formed at 7,750 generations, and the end state in (c) formed at 500 generations. In (b), the region around the PE Cell, located at grid cell number 447 and red in color, is active during all generations and has a much higher number of neighboring cells active than around the same location in (d).

doi.org/10.32388/25T20N

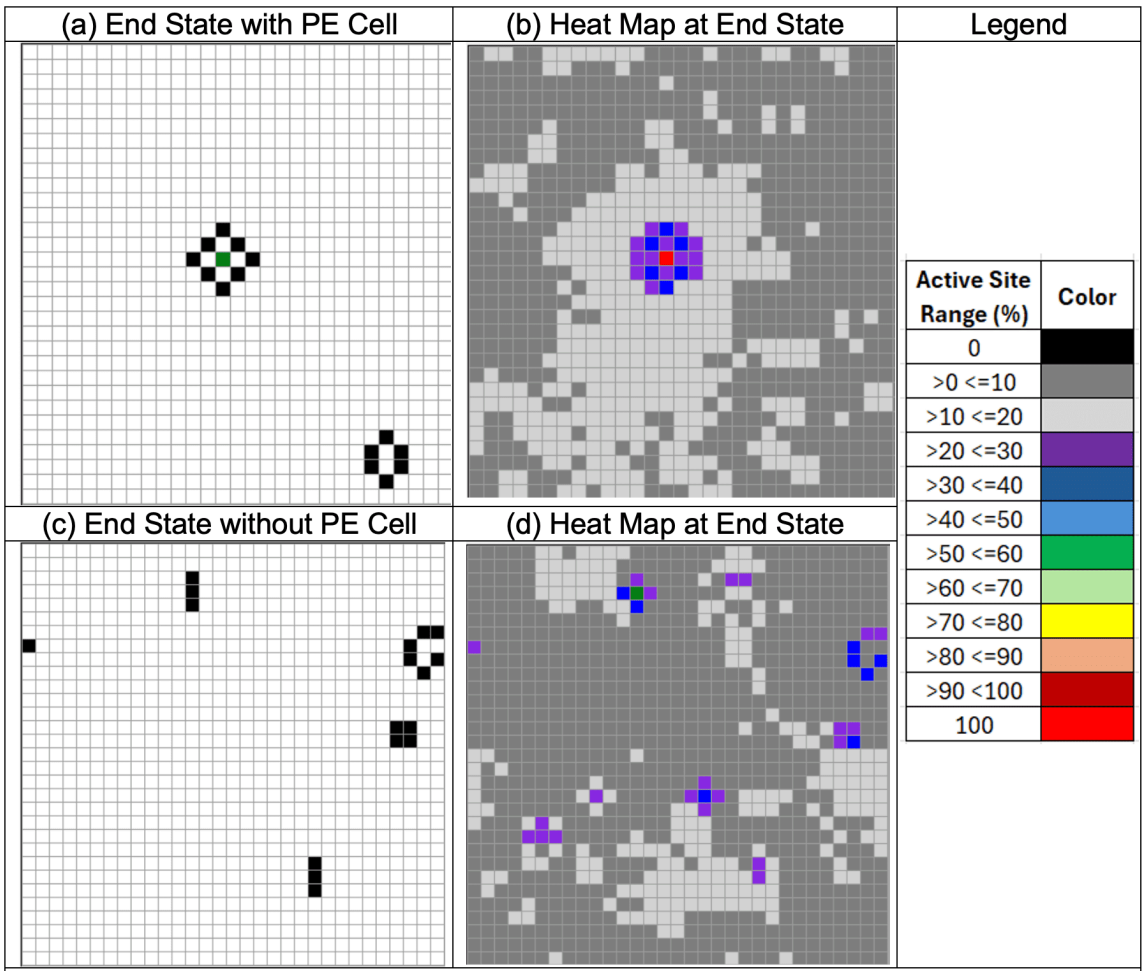


Figure A3.2. Heat Maps for the Diamond. The end states of evolution with and without the PE cell are shown respectively in (a) and (c). In (a), the Diamond was formed at 9,150 generations, and the end state in (b) formed at 400 generations. In (b), the region around the PE Cell, located at grid cell number 447 and red in color, is active during all generations and has a much higher number of neighboring cells active than around the same location in (d).

A.4. Quantifying Injection of Order

In my previous research[7], the quantification of injected order was measured by a Dissipation Factor (DF). Dissipation in this context is the power-law-driven behavior CA exhibit during the construction of static patterns, defined as the difference in the rate per generation of active sites becoming inactive versus inactive sites becoming active over the course of evolution. The difference in DF values for a soup evolving without the perennial seed and the same soup with the seed is termed a Differential Dissipation Factor

(DDF) and expressed as a percentage. For the phase 1 seeds, the variation in DDF for each seed is shown in Figure A4.1.

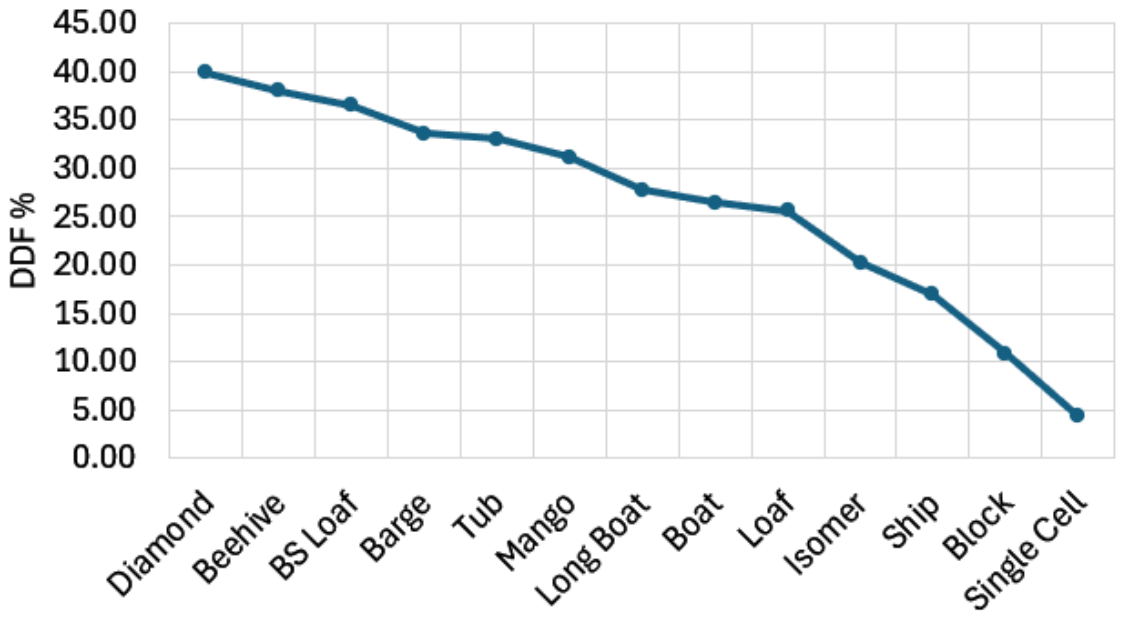


Figure A4.1. Differential Dissipation Factors. The effect of each phase 1 seed's geometric configuration of perennial cells results in a significant increase in the injection of order during evolution compared with phase 0's single perennial cell as a seed.

The degree of increase is approximately correlated with the number of open neighborhood binding locations that occur with diagonal cells on the periphery of a given seed's geometry[7]. For example, a Diamond has 8 diagonal cells on its periphery, greater than those of the other patterns in Figure A4.1, corresponding to their relatively smaller DDF% values. The effect of an increased injection of order was also discussed in Appendix A.3 relative to the dynamics surrounding a perennial seed's attraction of canonical cells.

Supplementaries

Patterns.xlsx

doi.org/10.32388/25T20N

Statements and Declarations

Funding

No specific funding was received for this work.

Potential Competing Interests

No potential competing interests to declare.

Data and Code Availability

The software and source code developed for this project are open-source and available for download via GitHub.

Use of Generative AI

Google's Gemini large language model was used to discover peer-reviewed resources for a portion of this research. Microsoft's CoPilot was used to discover peer-reviewed research regarding the Pareto probability distribution.

Acknowledgments

This paper is related to my Qeios paper "Functional Information of a Driven Cellular Automaton", a phenomenological study submitted to gather feedback. The feedback occurred and was helped, through 4 versions, by reviewers Dr. Al-Dabbagh, Serge Dolgikh, Fabrizio Browen, and Szymon Lukaszyk: thank you for the comments, insights, and analysis.

References

1. ^Rupe A, Crutchfield JP (2024). "On Principles of Emergent Organization." Phys Rep. **1071**:1–47. doi:10.1016/j.physrep.2024.04.001.
2. a, bWong ML, Cleland CE, Arend D Jr, Bartlett S, Cleaves HJ II, Demarest H, Prabhu A, Lunine JI, Hazen RM (2023). "On the Roles of Function and Selection in Evolving Systems." Proc Natl Acad Sci USA. **120**(43):e2310223120. doi:10.1073/pnas.2310223120.
3. ^Szostak JW (2003). "Functional Information: Molecular Messages." Nature. **423**(6941):689. doi:10.1038/423689a.

4. ^a, ^bHazen RM, Griffin PL, Carothers JM, Szostak JW (2007). "Functional Information and the Emergence of Biocomplexity." Proc Natl Acad Sci USA. **104**(suppl_1):8574–8581. doi:10.1073/pnas.0701744104.
5. ^a, ^bHazen RM, Wong ML (2024). "Open-Ended Versus Bounded Evolution: Mineral Evolution as a Case Study." PNAS Nexus. **3**(7):248. doi:10.1093/pnasnexus/pgae248.
6. ^Hazen RM, Wong ML (2026). Time's Second Arrow: Evolution, Order and a New Law of Nature. New York, NY: W. W. Norton & Co.
7. ^a, ^b, ^cRainwater JH (2024). "Self-Organization and Phase Transitions in Driven Cellular Automata." Artif Life. **30**(3):302–322. doi:10.1162/artl_a_00437.
8. ^Gardner M (1970). "Mathematical Games." Sci Am. **223**(4):120–123. doi:10.1038/scientificamerican1070-120.
9. ^Beer RD (2020). "An Investigation into the Origin of Autopoiesis." Artif Life. **26**(1):5–22. doi:10.1162/artl_a_00307.
10. ^Hazen RM (2009). "The Emergence of Patterning in Life's Origin and Evolution." Int J Dev Biol. **53**(5-6):683–692. https://www.researchgate.net/profile/Robert-Hazen/publication/26321936_The_emergence_of_patterning_in_life's_origin_and_evolution/links/00b7d531f32666b39d000000/The-emergence-of-patterning-in-lifes-origin-and-evolution.pdf.
11. ^Adami C (2016). "What is Information?" Phil Trans R Soc A. **374**(2063):20150230. doi:10.1098/rsta.2015.0230.
12. ^Gould SJ (1989). Wonderful Life: The Burgess Shale and the Nature of History. New York, NY: W. W. Norton & Co.
13. ^Bak P, Chen K, Creutz M (1989). "Self-Organized Criticality in the Game of Life." Nature. **342**(6251):780–782. https://www.nature.com/articles/342780a0.pdf.
14. ^Sornette D (2009). "Dragon-Kings, Black Swans and the Prediction of Crises." arXiv preprint. doi:10.48550/arXiv.0907.4290.
15. ^Von Hippel PT (2005). "Mean, Median, and Skew: Correcting a Textbook Rule." J Stat Educ. **13**(2). doi:10.1080/10691898.2005.11910556.
16. ^Hoehler TM, Davila A, Parenteau N, Quinn R, Des Marais D, Shkolyar S, Bebout L, Benner S, Blank JG, Brinckerhoff W, Cable M, Cady S, Eigenbrode J, Everroad R, Getty S, Glavin D, Gliniewicz S, Graham H, Hays L, Jahnke L, Lafuente B, Lau G, Lehmer O, Lima-Zaloumis J, Lynch K, Murphy A, Neveu M, Perl SM, Rainwater JH, Ricco A, Rios A, Som S, Wilhelm MB, Pohorille A (2025). "Introduction to the Life Detection Knowledge Base Project." Astrobiology. **25**(7):451–453. doi:10.1089/ast.2024.0107.
17. ^Shkolyar S, Bebout L, Blank JG, Cady SL, Cavalazzi B, Corbin E, Davila AF, Des Marais D, Fisk M, Hickman-Lewis K, Lima-Zaloumis J, McLoughlin N, Murphy AE, Noffke N, Perl SM, Pohorille A, Potter-McIntyre SL, Rainw

*ater JH, Westall F (2025). "Structural Biosignatures—A Category of Potential Biosignatures in the Life Detecti on Knowledge Base." Astrobiology. **25**(7):482–497. doi:10.1089/ast.2024.0105.*

18. ^*Sharma A, Czégel D, Lachmann M, Kempes CP, Walker SI, Cronin L (2023). "Assembly Theory Explains and Quantifies Selection and Evolution." Nature. **622**(7982):321–328. https://www.nature.com/articles/s41586-023-06600-9.*
19. ^*Patarroyo KY, Sharma A, Walker S, Cronin L (2023). "AssemblyCA: A Benchmark of Open-Endedness for Disc rete Cellular Automata." Second Agent Learning in Open-Endedness Workshop. https://openreview.net/foru m?id=5cEQ4ZOsIN.*
20. ^*Johnson JA, Fields BD, Thompson TA (2020). "The Origin of the Elements: A Century of Progress." Phil Trans R Soc A. **378**(2180):20190301. doi:10.1098/rsta.2019.0301.*
21. ^*Hazen R, Morison S (2020). "An Evolutionary System of Mineralogy. Part I: Stellar Mineralogy (>13 to 4.6 G a)." Am Mineral. **105**(5):627–651. doi:10.2138/am-2020-7173.*
22. ^*Schwartzman DW, Shore SN, Volk T, McMenamin M (1994). "Self-Organization of the Earth's Biosphere-Geo chemical or Geophysiological?" Origins Life Evol Biosphere. **24**(5):435–450. doi:10.1007/BF01582019.*
23. ^*Smith E, Morowitz HJ (2016). The Origin and Nature of Life on Earth: The Emergence of the Fourth Geosphe re. Cambridge University Press.*
24. ^*Pross A (2013). "The Evolutionary Origin of Biological Function and Complexity." J Mol Evol. **76**(4):185–191. h ttps://link.springer.com/article/10.1007/s00239-013-9556-1.*
25. ^*Pross A (2026). Life's Chemical Secret: The Origin of Purpose and Mind. Cambridge, Massachusetts: The MIT Press.*

Supplementary data: available at https://doi.org/10.32388/25T20N

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.