

Research Article

Functional Information of a Driven Cellular Automaton

Hank Rainwater¹

1. Independent researcher

For decades cellular automata (CA) simulations have proven versatile in modeling a wide range of complex natural phenomena due to their limited dimensionality and ability to allow adjustment of the small number of parameters that control self-organization. In this study, a model was designed that extends the self-organized phase of an automaton by keeping a single cell in a permanently (perennial) active state. After interacting with over 400,000 initial configurations of randomly placed standard automata cells surrounding the perennial cell, 351 unique and geometrically complex stable structures were created by the end state of evolution. These structures are analyzed with the goal of quantifying their complexity with Functional Information (FI) theory based on the degree of function of a structure within the CA environment using the probability of occurrence, quantified as information in bits. The expected selection bias inherent in CA is observed in this model and FI theory classifies each unique geometric structure according to its function, minimizing observer bias regarding complexity – frequently expressed as “I know it when I see it” – by supplementing intuition with information. In addition, a limited correlation was established between the median number of generations required to construct each structure and its probability of occurrence as measured by functional information theory. This link between probability of occurrence and construction in generations quantifies, within limits, the computational work performed by the state transition rule of the CA.

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

1. Introduction

Quantifying complexity created by the emergence of self-organized structures in nature tax the analytical tools of science. In a seminal review by Rupe & Crutchfield^[1] they comment that “After more than a century of concerted effort, physics still lacks basic principles of spontaneous self-organization.” Identifying the control parameters is often where the hard work begins in understanding these systems where self-organization emerges during phase transitions – think of liquid water becoming ice – as the system reaches a critical point.

This is why for decades cellular automata (CA) simulations have proven versatile in modeling a wide range of natural phenomena due to their limited dimensionality and ability to allow adjustment of the small number of parameters that control self-organization. Examples of CA applications include traffic flow^[2], fluid dynamics^[3], biology (Grander & Glazier, 1992), urban planning^[4], epidemiology^[5], self-reproduction^[6], artificial chemistry^[7], quantum mechanics^[8] and autopoiesis and multicellularity^[9]. With the increasingly important field of structural biology^[10], where the chemical structures of life are studied, CA can also contribute to this domain of knowledge.

In previous research^[11] I demonstrated how a modified version of Conway's Game of Life (GoL)^[12], called perennial evolution (PE), can create stable and increasingly complex geometric patterns in phases of evolution from random initial configurations. This evolution, while non-Darwinian, is characteristic of complex systems with the ability to "dissipate energy or to maximize entropy production through patterning"^[13]. PE is identified as *driven* because it continually injects order during evolution by keeping a CA cell or structure permanently active while maintaining and extending the system's dissipative action due to the physics of its state transition rule (STR)^[14]. The structures created in PE display a selection bias and a variable chemistry-like reactivity, acting as an artificial chemistry. When these structures, considered as artificial molecules, and their history of formation are analyzed by the methods of functional information (FI) theory^[15], a new understanding of self-organized complexity is obtained for this driven CA. The goal of this work is to gather Insights from the application of FI to CA to provide new opportunities for modeling the complexity of self-organizing systems.

2. The Perennial Evolution Model

As an elaboration on my previous work, where several of the frequently occurring GoL still life patterns were modeled as artificial molecules, the current project considers how to first build artificial molecules from a single perennial cell. The model in this research is analogous to how the primordial elements created by stellar evolution in the early universe^[16] eventually gave birth, over cosmological ages, to minerals that formed the geospheres of planets^[17] in continuing phases of structure building^[18].

2.1. Model Definition

As an extension of Conway's canonical GoL (CE), PE defines two types of active grid cells that can exist as shown in Figure 1, functioning as two-dimensional analogs of atoms that can bond and become structures representing *artificial* molecules. It is the construction process of this bonding and the resulting selection bias toward stable geometric patterns that is the focus of this research and not insights into the physics of creation of *real* molecules.

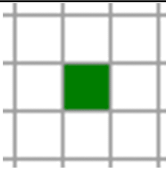
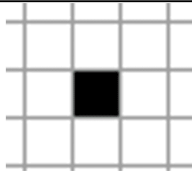
(a) Perennial Cell	(b) Canonical Cell
	

Figure 1. Atomic Components. Building blocks for construction of molecules in PE.

The bonding between the cells in Figure 1 is controlled by the physics of the STR that acts synchronously in discrete steps (*generations*) on all cells in the grid (*universe*) defined by 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 prior to the beginning of a phase of evolution. A perennial cell is permanently active throughout all generations and can have as many active canonical neighbors at the 8 cardinal and intercardinal directions of the cell (the Moore neighborhood) 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 molecular object if it becomes a static geometric pattern during evolution and remains static until the universe reaches a quiescent end state. This end state occurs when the density of active cells in the evolving automaton reaches a constant or periodic value, but the pattern itself must be static. Also, the 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 life will be created in PE because rules 1 and 2 are identical with the CE rules. However, patterns will also be created that cannot exist without a single perennial cell and these objects will have a special focus in the analysis of this research and referred to as PE molecules. An example of PE construction is shown in Figure 2 where the creation of still life objects occurs from a random initial collection of active CE cells and a single PE cell.

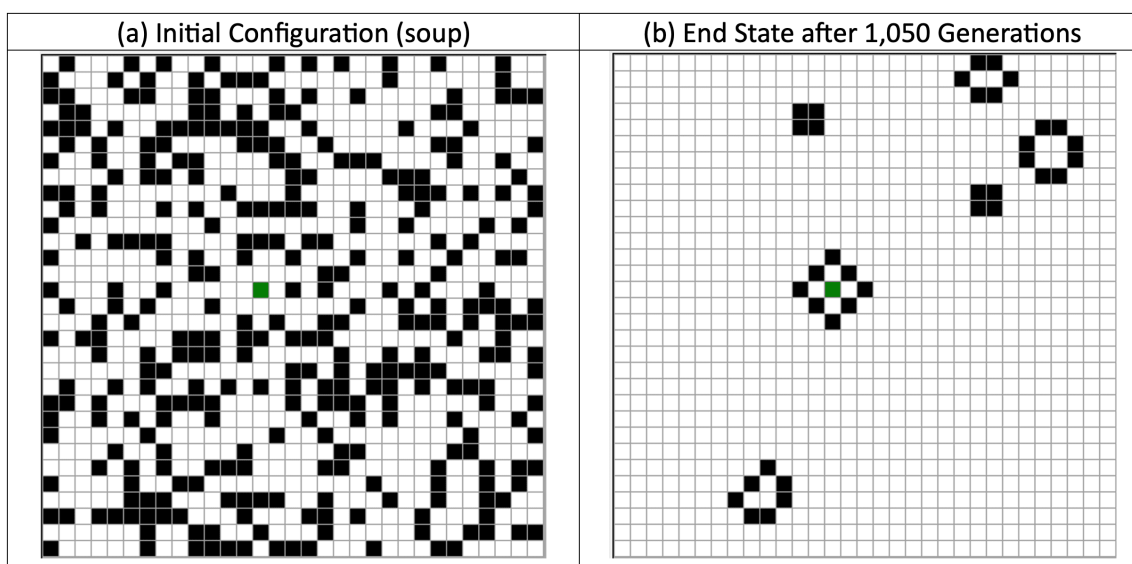


Figure 2. Perennial Evolution. A single perennial atom within a random soup of canonical atoms in (a) evolves after 1,050 generations to a static density in (b) creating a stable structure modeled as a molecule that could not exist without the perennial cell. This PE molecule at the center of the grid in (b) is named a 'Diamond'. Several other classic CE still life also appear at the end state. The universe size for this simulation is a 31 X 31 toroidal grid with an initial density of active cells in (a) of 35%.

With few exceptions it is common in PE for an embedded perennial cell to extend the number of generations for a random initial configuration of canonical cells to reach a static end state as compared with the same soup without the perennial cell. As an artificial chemistry this phenomenon can be viewed as the reactivity of the perennial cell – a type of catalytic effect – with the canonical cells. For example, in Figure 2, the soups in (a) without the perennial cell reached a static end state in 180 generations as compared to the 1,050 when a perennial cell is present. This catalytic effect – an extension of the region of dissipation during evolution where the rates of change per generation between active cells becoming inactive (death) and inactive cells becoming active (birth) are competing – is the region of self-organization that exhibits power law behavior^[19]. These death/birth rates, called dissipation factors can be measured and correlated with the number of generations that a soup requires to reach a static end state and demonstrate a power law correlation and are discussed in my previous work on perennial evolution^[11].

To emphasize the fact that PE created molecules are a *simulation* of a chemistry process, from this point forward molecules will be referred to as two-dimensional *patterns* that have a stable structure at the end state of evolution. The atoms will also be referred to as active *cells* to further clarify the nature of the model whose

primary purpose is to demonstrate self-organized complexity in a system of many locally interacting components.

2.2. Populating the Model with Data

Statistically significant ensembles of constructed patterns in PE are only revealed from evolution of a large set of initial state configurations (soups). Each soup contains a single perennial cell embedded among randomly positioned canonical cells. Each soup is unique and the position in the grid of a canonical cell is assigned by a random number generator each time a soup is created by the software designed for this project, but the PE cell is assigned to the same grid location for all random soups. A toroidal grid size of 31 rows and 31 columns with 961 cells for each soup has an active cell density of 35%, chosen because GoL dynamics are scale-invariant relative to grid size (Pena & Sayama, 2021, p.108) and this density is in the range where self-organized criticality is optimal^{[20][21]}.

The total number of initial soups in the dataset for this project was 416,000 and each soup was set to evolve up to a maximum of 2,000,000 generations. Most of these soups (401,275) reached an equilibrium state prior to the maximum generation limit and those that did not reach equilibrium or were periodic were discarded from analysis. The initial soups were divided into 300 subsets to allow computation on a cluster of desktop computers.

The expected results of the dataset were the creation of patterns but 42.79% of the 401,275 soups did not produce any structure: the perennial cell remained unbound to canonical cells. However, 229,570 of the initial configurations, or 57.21%, produced a total of 351 stable still life with 186 qualified as PE patterns, unique from CE patterns. The software developed for this project creates a signature for each pattern that defines its unique geometry by the number of cells per pattern as well as the number of bonds between the perennial cell and canonical cells. The patterns, as they were discovered during data population, are organized into families based on the number of canonical cells per pattern. In Figure 3 the number of patterns discovered per subset of data execution is presented, organized into families.

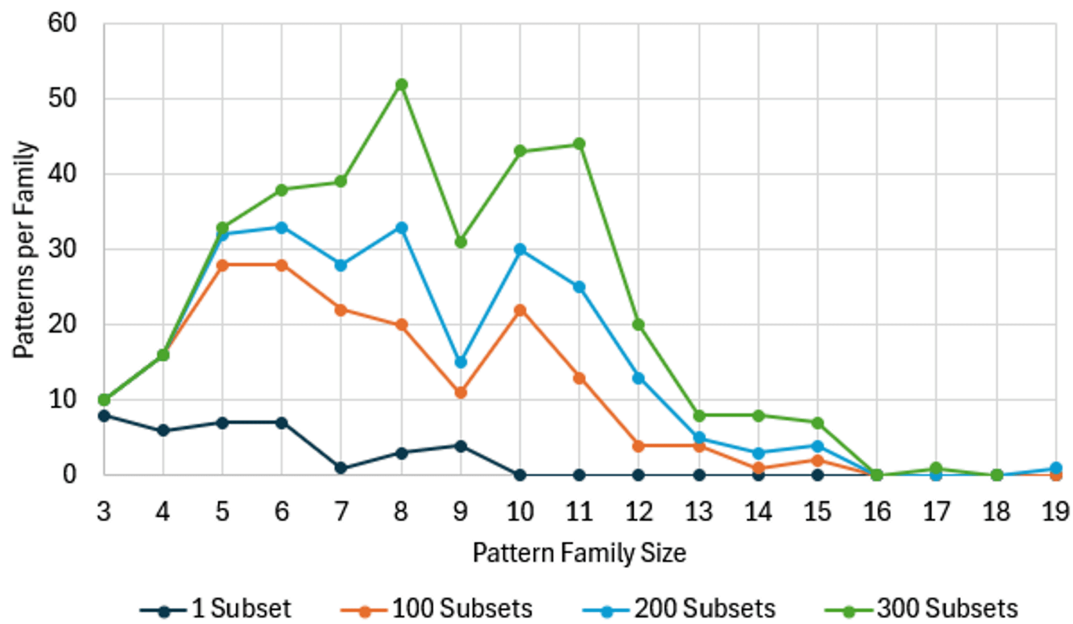


Figure 3. Discovery of Patterns. The first subset discovered 36 patterns, and the subsequent number discovered increased as additional subsets of soups were executed. The total number of subsets was 300, creating 351 total patterns. No patterns were discovered for family sizes 16 or 18 but are shown in the graph to highlight this fact.

The details of family patterns are discussed in the next section.

3. Pattern Phenomenology

The organizational scheme of patterns into families, based on the number of canonical cells per structure, was developed as limits to the variety of patterns that can exist within a family of a defined size were discovered. The smallest family size was 3 canonical cells bound to the single perennial cell and the largest was 19. A view of family statistics is shown in Figure 4.

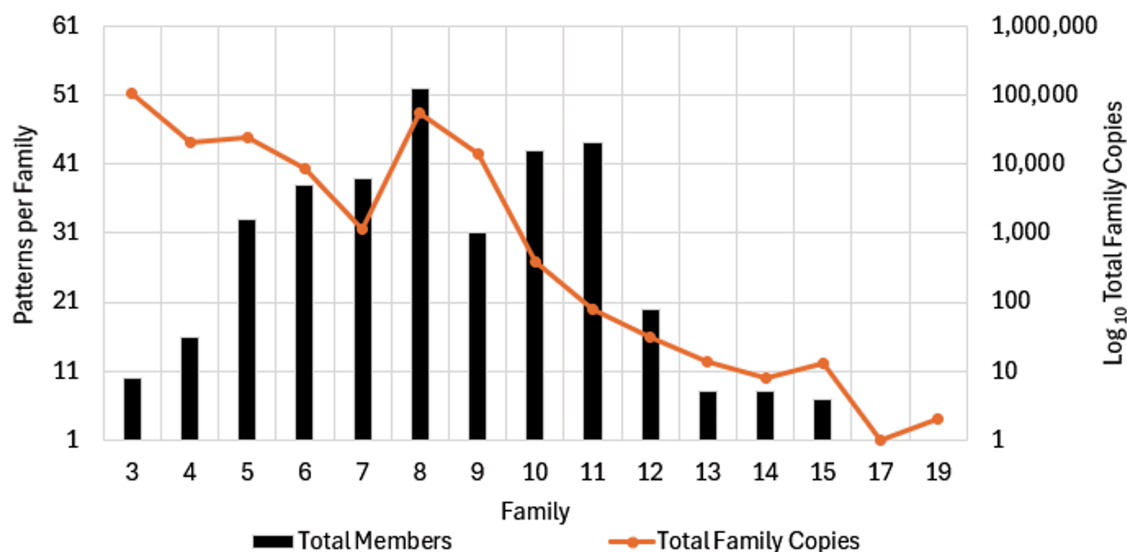


Figure 4. Family Statistics. Family 8 contains the PE pattern named ‘Diamond’ shown in Figure 2 (b) with over 53,000 copies Any correlation between the number of pattern copies per family versus the family size is not apparent in this data although there is a general falloff after family 9.

3.1. Minimum Pattern Size Limit

Patterns of less than three canonical cells cannot exist in PE nor in CE as illustrated in Figure 5.

	(a)	(b)	(c)	(d)	(e)
Initial State					
End State					

Figure 5. Minimum Pattern Size. In (a) the end state is not a valid still life pattern. In (b) this structure becomes the classic period 1 oscillator known as a Blinker in CE and is thus not stable nor qualifies as a pattern. In (c) a stable Block is formed and in (d) a stable Tub. The end state of (a) through (d) occurred in a single generation while in (e) the end state was reached in 1,278 generations from evolution with a random soup of canonical cells in which the perennial cell was embedded; the initial state shown is generation 1,277. (The Block and Tub were also created from random canonical soups with a single embedded perennial cell but shown here in a simplified form of creation.)

The Block and Tub are part of a family identified as P_1C_3 since they contain 3 canonical cells bound to the single perennial cell, but they do not qualify as PE patterns since they can be created in CE. More on the PE pattern (e) in the next section.

3.2. Isomers in Families of Patterns

In addition to the minimum size of a pattern, many patterns in a family can appear in several orientations relative to the grid as two-dimensional spatial isomers (*stereoisomers*). For example, family P_1C_3 contains 8 stereoisomeric members as shown in Figure 6.

Pattern								
Copies Created	217	211	209	204	202	201	197	181
Bond Location	NE	SW	SW	NE	SE	NW	SE	NW

Figure 6. Family P_1C_3 Stereoisomers. The perennial cell's bond location to a canonical cell varies among these patterns corresponding to the intercardinal directions of the Moore neighborhood. Rotation and/or reflection with respect to the XY-plane of any one of these patterns demonstrates the structure's isometry. The total number of copies for all forms of this pattern is 1,622. This geometric form is unique to PE since it can't exist without the perennial cell.

Stereoisomers reveal that for a given family, e.g., P_1C_3 , there is a limit to the number of unique patterns that can exist within a family of a specific number of canonical cells. If a pattern like those in Figure 6 is manually assembled where the first canonical cell is bound to the perennial cell's north binding location, the structure is not stable as seen in Figure 7.

Initial State	Generation 1	Generation 2	Generation 3		End State at 168 Generations
					

Figure 7. Nonstable Hypothetical P_1C_3 Object. A binding to the north of the perennial cell – as well as a binding to the east, south, or west – in the absence of any other canonical cells in the universe, results in the initial state evolving to a static configuration in 168 generations, where a stable Diamond pattern is created.

The behavior of the evolution shown in Figure 7 results from the driven nature of PE where the presence of a perennial cell with only a few canonical cells injects order as the universe evolves. In this context the single PE cell acts as a catalysis within the CA universe of potential canonical cells, reacting at a rate governed by the STR. The position of the canonical cells bonded to the perennial cell in Figure 7 illustrates how critical geometry is to the final state of such a pattern. This sensitivity to position, a result of the STR that governs cell bonding, is demonstrated by the possible number of isometric forms that can exist for other families as seen in Figure 8.

Family Pattern	P_1C_4		P_1C_5			P_1C_6
						
PE Pattern?	True	True	False	True	True	False
Forms Detected	4	8	4	8	8	8

Figure 8. Isometric Patterns. In family P_1C_4 the two isomers complete the possible number of structures that can exist in this family (the Boat, discussed shortly, in its 4 forms are the remaining patterns in this family). The repeated appearance of 4 and 8 forms of structures is a function of the CA neighborhood with its 8 possible directions of bonding.

3.3. Selection Bias

Of the 351 patterns created by the dataset Figure 9 shows those patterns exhibiting a high degree of selectivity and are ranked from left-to-right by their number of copies.

Family	3	8	3	5	4	5	9	6	5
Pattern									
Name	Block	Diamond	Tub	Beehive	Boat	Barge	Yacht*	Loaf	Ship
Copies	94,219	53,337	10,696	5,839	5,055	4,460	3,480	2,042	1,855
Other Forms and Copies				 5,835	 5,031	 4,257	 3,474	 2,038	 1,853
				 5,010		 3,432	 1,979		
				 4,960		 3,425	 1,963		

Figure 9. Frequently Occurring Patterns. These patterns will be a primary focus for analysis and represent 98% of the 229,570 total patterns created from the dataset. See Supplementary Data (All Patterns.xlsx) for all 351 molecules in a spreadsheet format. *The ‘Yacht’, whose GoL name is ‘Barge Siamese Loaf’, is not used because a shorter monicker was needed for ease of reference.

PE did produce many patterns within the top 10 frequently occurring GoL still life though in a slightly different order relative to their GoL rank, however the GoL rank (see [LifeWiki](#)) is based on CA evolution with a different grid size, style and initial density so that a direct comparison cannot be made with the PE rank.

4. Functional Information Theory for Perennial Evolution

The selection bias demonstrated in Figure 9 is a result of the deterministic STR rules acting on each cell in the PE universe with each of its eight neighbors at each discrete generation of the 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”^[22]. In this view of biology, evolution is driven by random genetic mutations with natural selection acting as a *filter* allowing reproductively successful mutations to survive. In the context of a driven CA, the single PE cell also functions as a type of filter by acting as an *attractor* of canonical cells, forming 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 PE cell while the dynamics in the rest of the universe follow the canonical GoL rules as stated in the model definition (see Appendix A1 for an example). This attractive *function* of the PE cell is defined in FI theory as a “degree of

function (E_X) [that is] a quantitative measure of a configuration's ability to perform the function X "^[23]. This function is the system's ability to create stable patterns, and the "quantitative measure" is given by the formula:

$$I(E_X) = -\log_2[F(E_X)] \quad (1)$$

In the context of PE, the "degree of function (E_X)" is a measure of the action performed by the STR to form a stable pattern^[24] with ≥ 3 active canonical cells with a single perennial cell as demonstrated above in Figure 5. The $F(E_X)$ term in Equation 1 is the "fraction of all possible configurations of the system that achieves a degree of function $\geq E_X$ " and is defined in this research as the Functional Information Detection Ratio (FIDR) as follows...

$$\text{FIDR} = N_{\text{copies}} / N_{\text{soups}} \quad (2)$$

where...

- N_{copies} = the number of copies of a specific pattern; this number also represents the number of soups that created the copies since each soup results in one copy,
- N_{soups} = 401,275 i.e., the total number of soups producing all patterns (§ 2.2).

FIDR thus defines a measure of the probability of a pattern's occurrence within the evolving system so that Equation 1 can also be written as $FI = -\log_2[\text{FIDR}]$. As a result of this formal definition, the selection bias, an aspect of complexity observed in CA, is encapsulated via the FIDR factor. The concept that a function can be equated with the stability of a geometric object has proven useful in other work^[25] and will prove its merit in this study with CA.

The 401,275 configurations (from the initial 300 subsets of soups) are the "all possible configurations" where each configuration could result in the creation of a pattern, even though only 229,570 configurations did create a pattern and achieve the defined degree of function. However, FI considers the larger value of 401,275 soups as representative of the state of knowledge about the system prior to any evolutions of the automaton, that is, all configurations are initially equally likely to produce patterns of the defined degree of function because the probability distribution of the system is uniform – and unknown – prior to any evolutions being performed^[26]. Once evolution begins, information about the system increases with the value of $I(E_X)$ in bits quantifying this new information. It is also important to note that the functional information of a pattern is not the property of a pattern in isolation, "but of the ensemble of all possible sequences, ranked by activity"^{[15][27]}.

In Figure 10 FI theory is applied to the frequently occurring patterns shown in Figure 9 of the previous section and shows the graphically inverse relationship between FIDR and FI values.

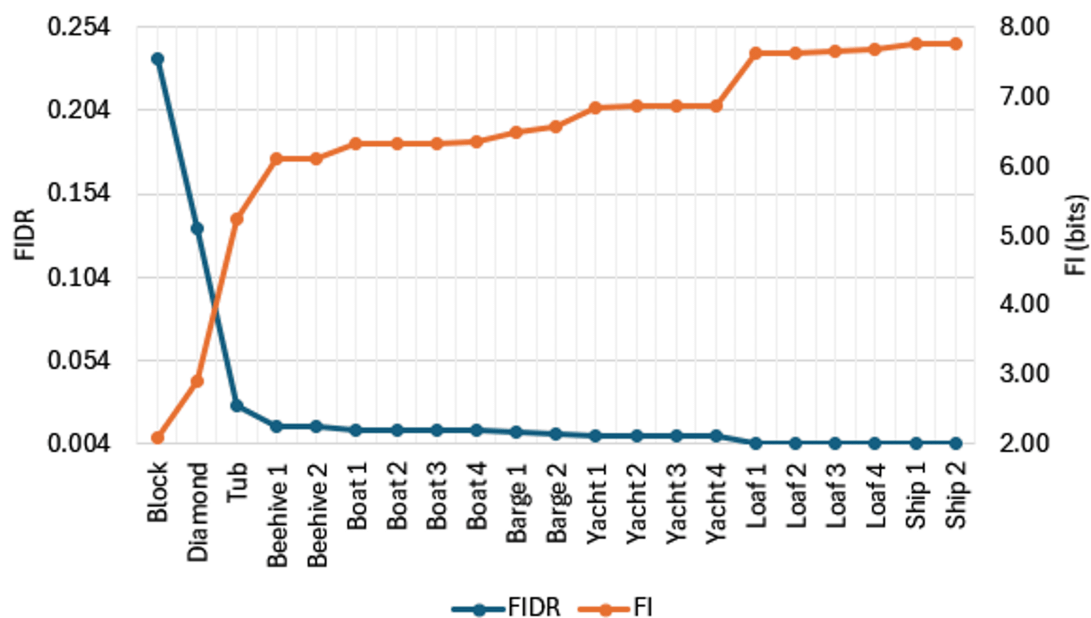


Figure 10. FIDR vs FI for Frequently Occurring Patterns. The Block, a CE pattern, and the Diamond, a PE pattern, are the two most frequently occurring patterns.

All 351 patterns are shown in the next graph and contain a reference to several patterns from Figure 10 to enable a perspective on the less frequently occurring patterns.

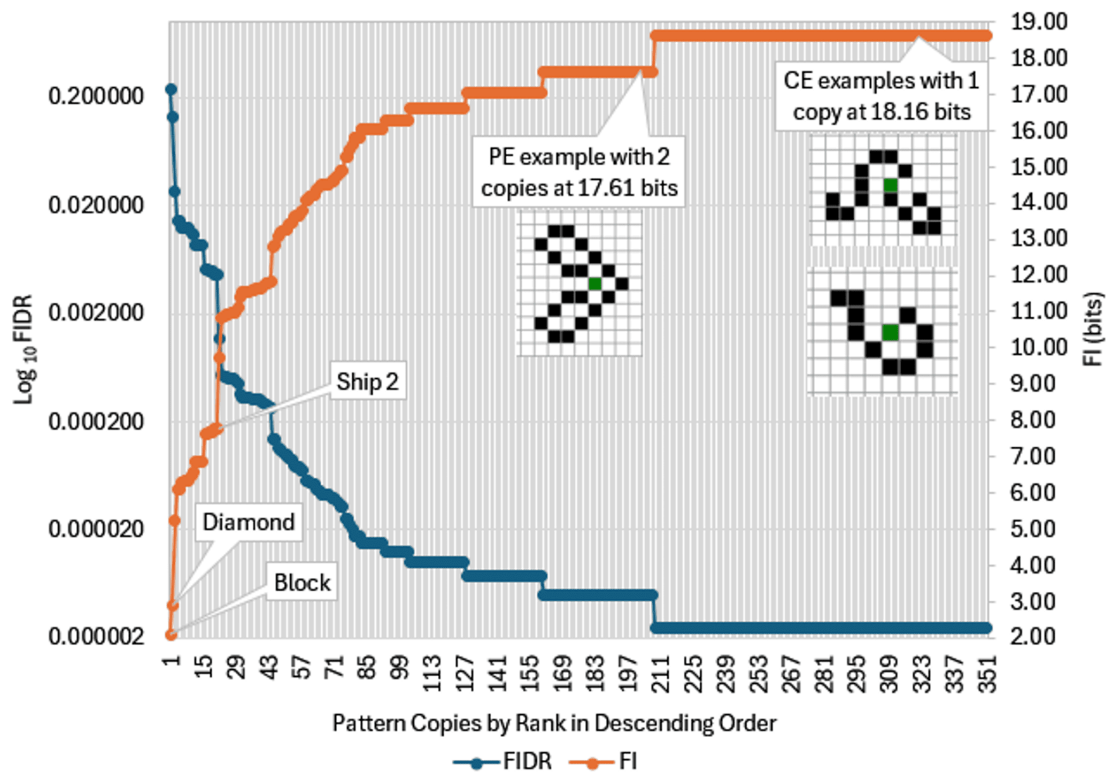


Figure 11. FIDR vs FI all 351 Patterns. The true shape of the FIDR plot is accurately shown in the previous plot, Figure 10, as L-shaped but here the FIDR axis is shown on a logarithmic scale to highlight the range of probabilities that rapidly decrease from the Block at a FIDR of 0.23 and a FI of 2.09 bits.

While the images of infrequent patterns in Figure 11 offer a picture of geometric complexity that can be evaluated along the lines of the time-worn saying, “I know it when I see it,” the degree of function quantified by FI supplements visual intuition with information. However, this information – FIDR and the equivalent FI value in bits – is the basis of the theory’s definition of complexity and does not explicitly include information related to how the geometric structure of patterns with identical FI values differ such as two patterns with a single copy at 18.61 bits. In the next section FI theory will be used in relationship to the statistics of pattern construction to derive additional insights into complexity.

5. Analysis of Pattern Construction Using FI Theory

FI theory, as discussed above, reveals aspects of complexity related to frequency of occurrence but does not directly address the construction process that occurs during PE evolution. Topics concerning this process are investigated in this section.

5.1. Construction and Frequency of Occurrence

Frequently occurring patterns such as the Block and Diamond have a large FIDR value compared with the less frequent patterns shown above. Is it possible that these frequent patterns are more numerous because they are easy to construct? And does ‘easy to construct’ imply that the quantity of computation work performed during the soup’s evolution by the STR is less than that for the infrequent patterns that are harder to construct? Stated in terms of a hypothesis, does a more frequent pattern, as defined by FI theory, require fewer generations to form during evolution than a less frequent pattern?

To test this hypothesis, the software created for this project was designed to record the minimum, maximum, average, 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 (min, max, average and median) 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. For example, the Block, the most frequently occurring pattern, has a range of CS values as shown in Figure 12.

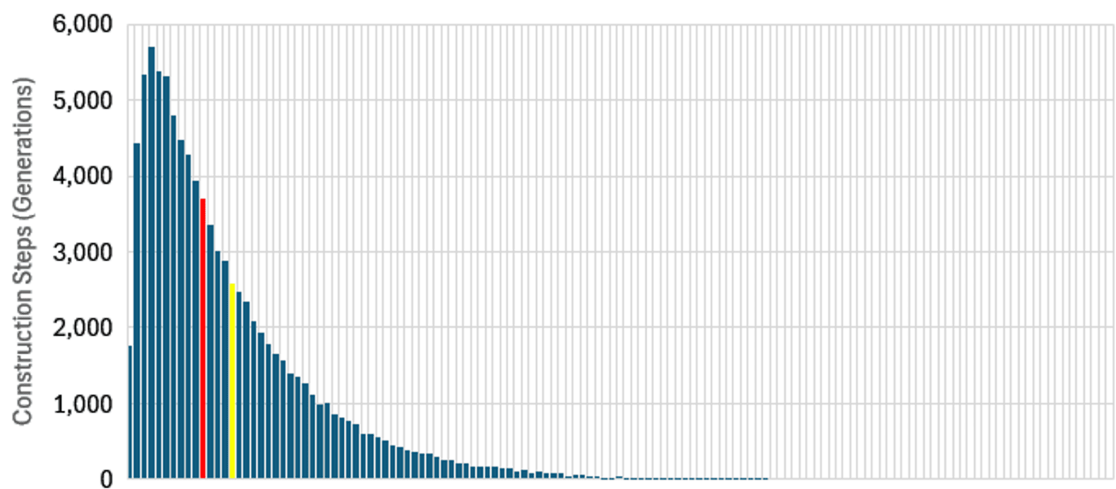


Figure 12. The Block’s Construction Steps. The minimum CS is 4 generations and the maximum 7,176. The average of 749 is shown by the yellow column and the median at 546 in red. The standard deviation of the average CS values is 678 generations. At each CS value, the Block is formed and remains stable until the end state of the soup is reached. The histogram’s horizontal axis contains 52 bins, but the bin ranges are not shown due to space limitations on the page.

The range of values in this histogram with its highly skewed long tail is common for CA evolution where a power law is active during the self-organizing phase. The shape is like those seen with a Pareto distribution^[28], and demonstrates black swan events^[29] in the long tail, that while less numerous, distorts the validity of the

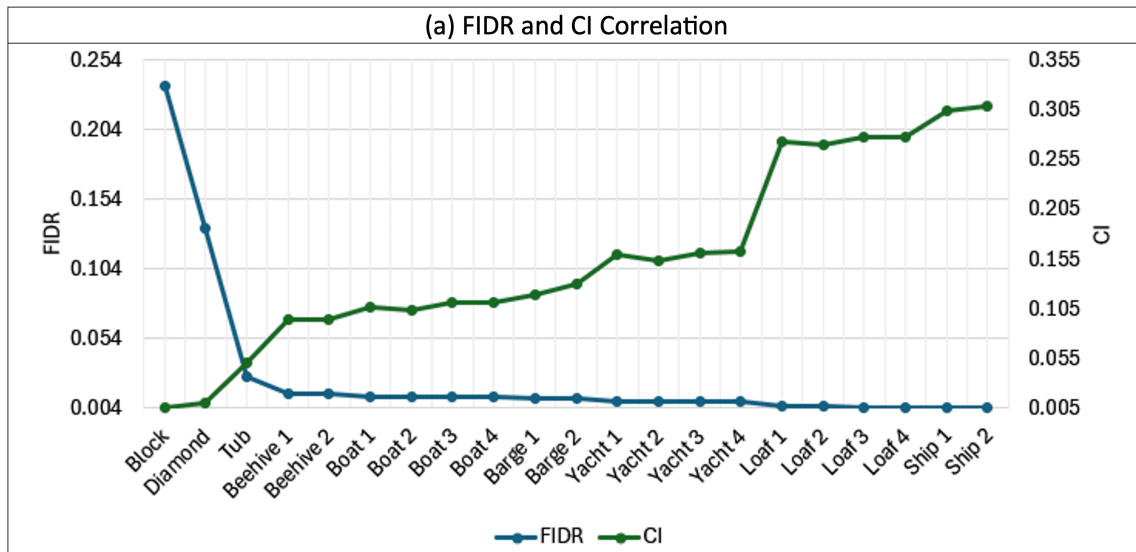
average as a *typical* value. For this reason, the median value is more representative of a typical value (see Appendix A2 for more details on the choice of using a median vs an average value for a CS metric). The other 20 frequently occurring patterns of Figure 9 have similarly skewed histograms of CS values representing the contingent nature of CA evolution where each value represents an instance of a specific pattern's construction in 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 construction generations per pattern copy with a dimension shared with the FIDR. 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 FIDR value from Equation 2 is:

$$\text{FIDR}_{\text{Block}} = \text{Block}_{\text{copies}} / N_{\text{soups}} = 94,219 / 401,275 = 0.2348 \text{ copies/soup}$$

The normalized CI value is calculated in a similar fashion:

$$\text{CI}_{\text{Block}} = \text{Median CS}_{\text{Block}} / \text{Block}_{\text{copies}} = 546 / 94,219 = 0.005795 \text{ generations/copy}$$

When the normalized CI values are compared with their equivalent FIDR values for the frequently occurring Figure 9 patterns their correlation is displayed in Figure 13.



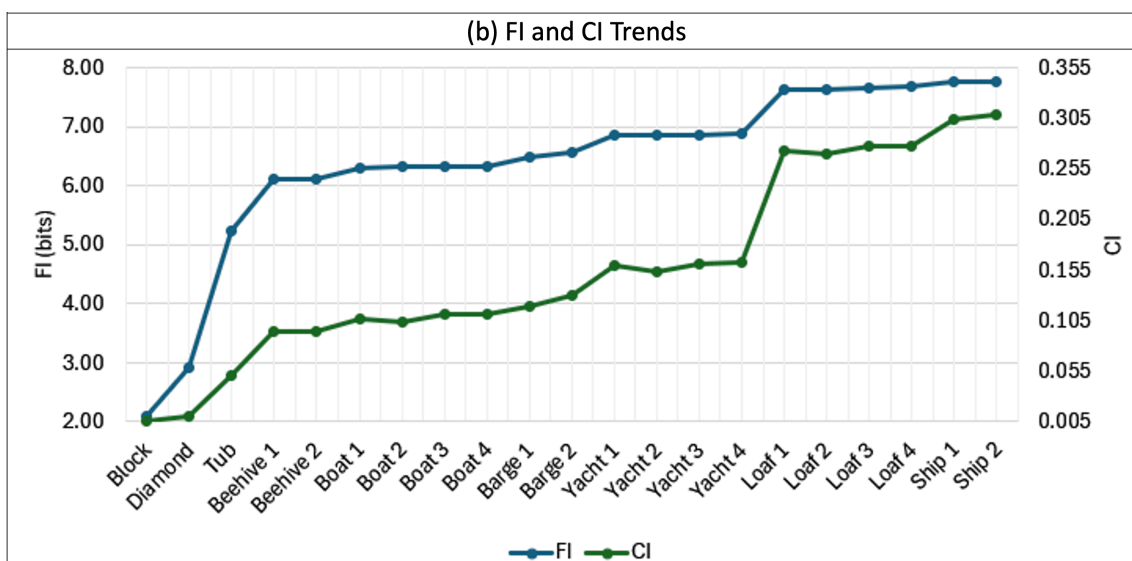


Figure 13. Construction Index Correlation. In (a) as a pattern's FIDR decreases its CI increases; in (b) while the values are incommensurate, the trends are similar. The most frequent patterns, the Block and Diamond, require fewer generations to create than the less frequent. The range of the number of copies per pattern in this graph is from 94,219 for the Block to 1,853 for Ship 2.

Based on the metrics of FI theory, the Block is less complex than Ship 2 where complexity is defined as the degree of function measured by frequency of occurrence. Using the CI metric, new information is added to the definition of complexity, namely that a Block is *typically* easier to construct than a Ship. Note that 'typically' is important relative to the CI metric because in the case of a Block vs a Ship, some soup configurations may create a Ship in fewer generations than other soups create a Block. For example, as recorded in Supplemental Data, the Block's 94,219 soups had a minimum CS of 4 generations with a maximum of 7,176 while Ship 2's CS ranged from 15 to 6,157 generations. Thus, this CI/FIDR correlation is based on the *total* ensemble of initial random soups and is a measure of the *typical* work done by the STR to create a particular pattern.

As the number of copies of a pattern decreases, the correlation shown above begins to degrade as shown in Figure 14.

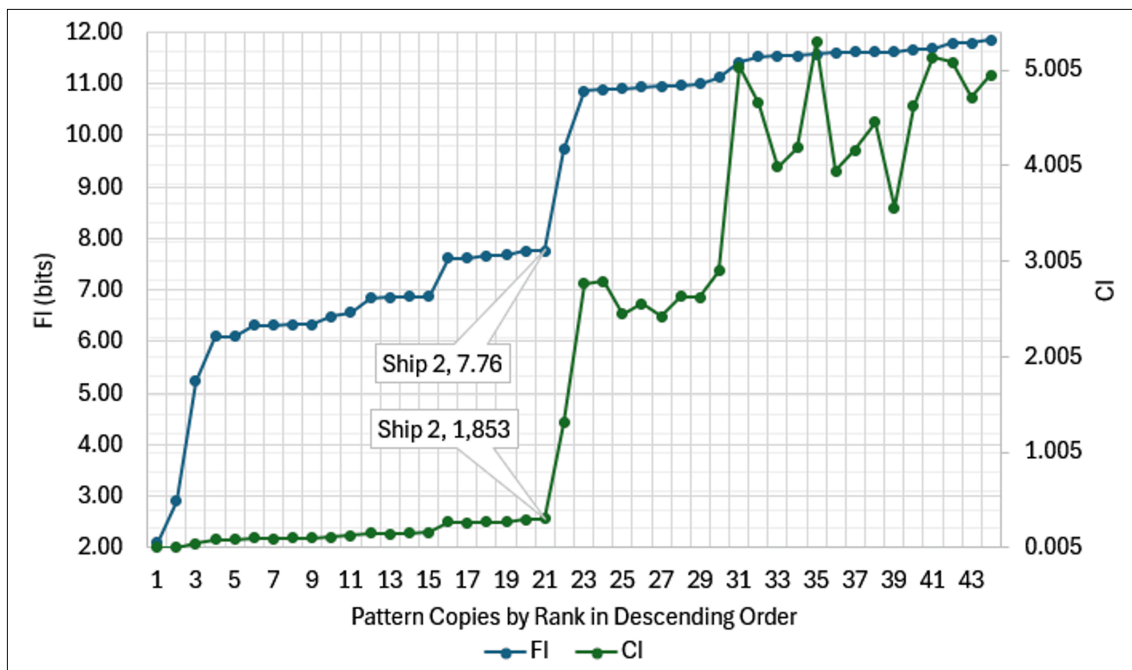


Figure 14. CI and FI Trends for 44 Patterns with Greater than 100 Copies. While the CI trend does correspond to increased construction effort for less frequent patterns, as the number of copies of a pattern decreases, the ‘noise’ in the CI trend increases. This will be discussed in detail below.

The term ‘noise’ is shorthand for the variability in median CS values determined from the skewed probability distribution illustrated for the Block in Figure 12, common for all patterns. However, with a sufficiently large number of copies for a given pattern, the noise is minimal resulting in the higher correlation between CI and FI trends in Figure 13 as compared with Figure 14. As noise increases, correlation becomes degraded and Figure 15 graphically illustrates the noise in the median CS values from which the CI is calculated.

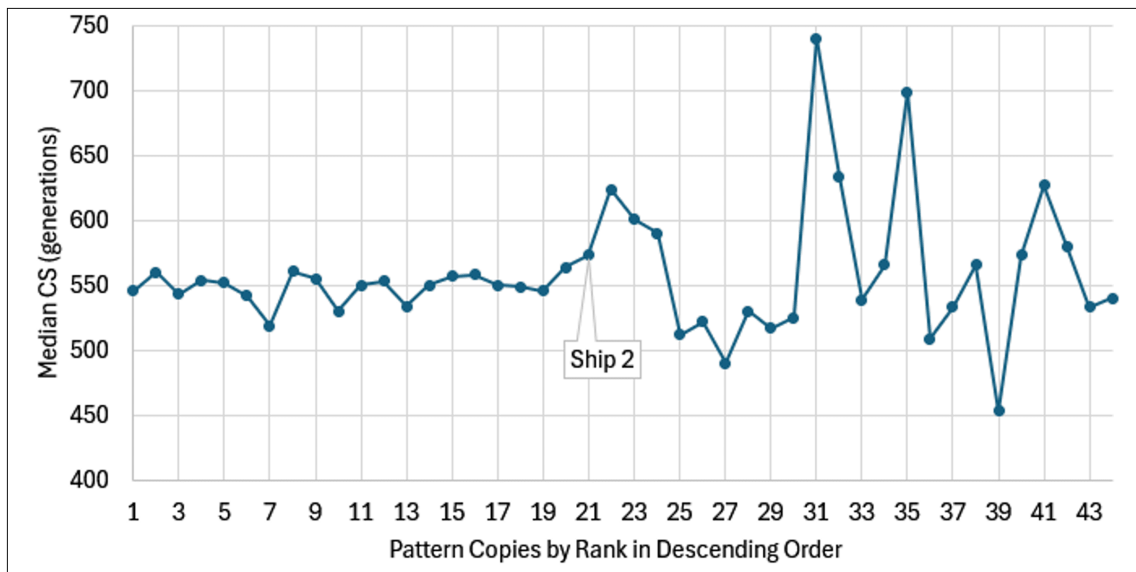


Figure 15. Increase in Median CS Variability with Decrease in Pattern Copies. The 44 patterns are shown here with Ship 2 at 1,853 copies as a reference to Figures 13 & 14. The last datapoint (44) in this graph represents a pattern with 109 copies.

A quantification of the Median CS variability is shown in Table 1.

Pattern Rank	Number of Copies	Average Median CS
1 to 22	Greater than 1,000	550 ± 12
23 to 44	100 to less than 1,000	565 ± 65
45 to 77	10 to less than 100	532 ± 141
78 to 200	2 to less than 10	678 ± 381
201 to 351	1 copy	796 ± 616

Table 1. Increase in Median CS Standard Deviation with decrease in Pattern Copies. The average median CS values include CE and PE pattern types for the collection of patterns of a given rank range.

In general, the noise increases as the number of pattern copies decreases. Table 1 demonstrates that the low median variability of patterns with greater than 1,000 copies resulted in the less noisy correlations seen in Figure 13 vs those in Figure 14 where the number of copies decreased. Also, the patterns with a single copy have

only a single CS value and thus any CI value is meaningless since their frequency of occurrence (FIDR) is constant at 0.000002. (See Appendix A2 for a comment on the low variability for those patterns with greater than 1,000 copies.)

5.2. Construction and Perennial Cell Binding

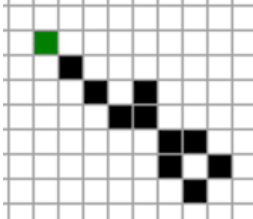
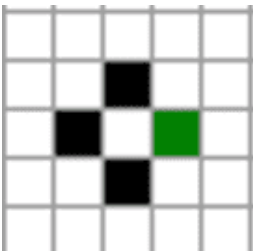
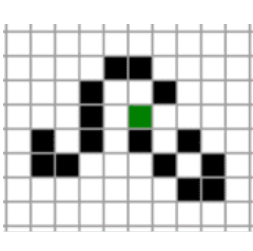
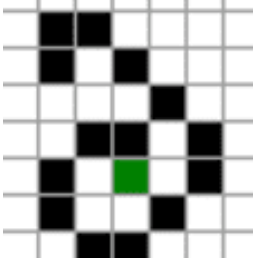
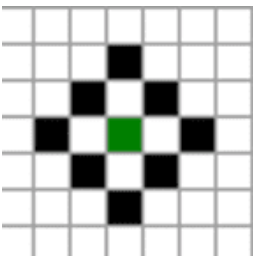
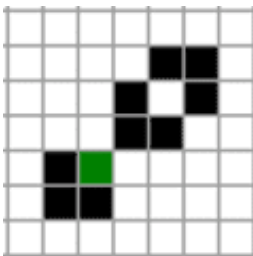
The various geometric structures of patterns and their relationship to the binding strength of the PE cell, and any correlation between binding and functional information, are considered in this section. In this context, *strength* is a function of the number of bonds the single perennial cell has to canonical cells. This binding is of course an analogy for valency, and due to the STR where the perennial cell can have as many active canonical neighbor cells as the B3/S23 rule will allow. The bond values range from 1 to 6 canonical cells bound to the single perennial cell and are shown segregated by pattern type per family in Table 2.

Family	CE Members	CE Bond 2	CE Bond 3		PE Members	PE Bond 1	PE Bond 4	PE Bond 5	PE Bond 6
3	2	1	1		8	8	0	0	0
4	4	0	4		12	8	4	0	0
5	10	2	8		23	23	0	0	0
6	20	14	6		18	18	0	0	0
7	33	19	14		6	6	0	0	0
8	26	15	11		26	4	17	5	0
9	20	4	16		11	2	8	1	0
10	19	8	11		24	2	18	4	0
11	15	5	10		29	8	13	8	0
12	5	1	4		15	0	13	1	1
13	7	4	3		1	0	0	1	0
14	3	2	1		5	0	4	1	0
15	1	1	0		6	0	6	0	0
17	0	0	0		1	0	1	0	0

Family	CE Members	CE Bond 2	CE Bond 3		PE Members	PE Bond 1	PE Bond 4	PE Bond 5	PE Bond 6
19	0	0	0		1	0	1	0	0
Totals	165	76	89		186	79	85	21	1

Table 2. Bond Values per Family. No patterns were detected for family sizes 16 or 18.

Several examples of the effect of the number of bonds on a pattern's geometric configuration and associated FI values are presented in Table 3.

PE Cell Bonds	Pattern Type	Number of Patterns	Range of FI Values in Bits	Most Frequently Occurring Example	Least Frequently Occurring Example
1	PE	79	10.85 to 18.61		
2	CE	78	5.23 to 18.61		
3	CE	87	2.09 to 18.61		
4	PE	85	2.91 to 18.61		

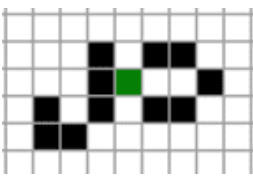
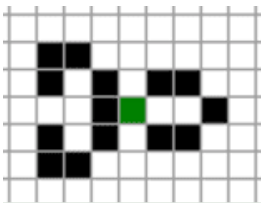
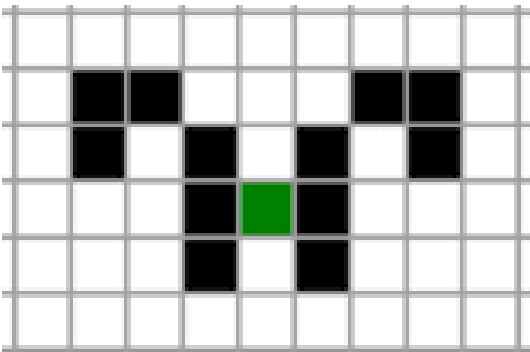
PE Cell Bonds	Pattern Type	Number of Patterns	Range of FI Values in Bits	Most Frequently Occurring Example	Least Frequently Occurring Example
5	PE	21	17.03 to 18.61		
6	PE	1	18.61		

Table 3. Functional Information and Cell Bonds. All bond values produced patterns with the largest FI values, but the bond values 5 & 6 only were able to produce rare PE patterns compared with the lower bond values. The bond value 6 only occurs for one rare pattern.

The PE patterns with 1, 4, 5 or 6 bonds could not be created by the GoL and thus the perennial cell's binding strength *may* be a factor affecting their frequency of occurrence. However, the CE patterns with 2 or 3 bonds include the frequently occurring patterns from Figure 9; the PE Diamond with 4 bonds is also part of this collection of patterns. From the data in Table 3, all bond values are represented in the rarest patterns at 18.61 bits occurring for a single pattern copy. Could family size be a factor in selectivity in addition to the bond type? To answer this question Figure 16 shows all 351 patterns plotted vs family size and FI value for context and to aid analysis.

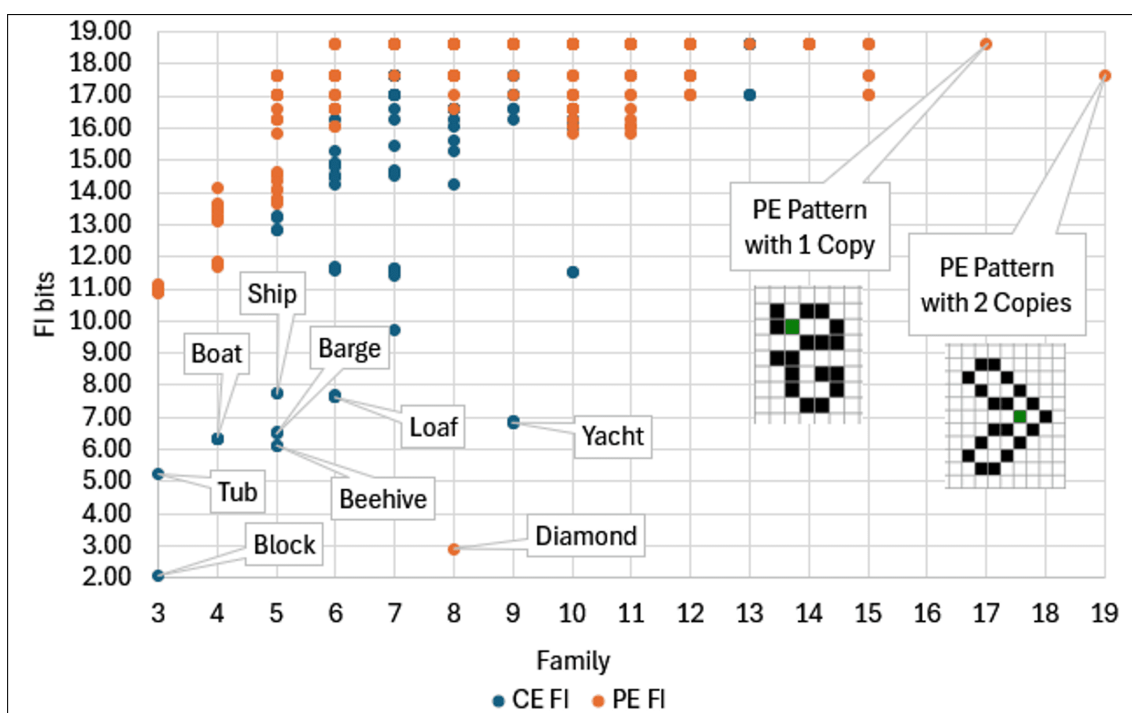


Figure 16. Selectivity and Families. The CE families contain 165 patterns and PE 186 but FI datapoints overlap in these plots so all 351 patterns are not visible. The patterns from Figure 9 with the largest selection bias are labeled by name and, for those patterns with more than one form, only a single instance name is shown. No patterns were detected for family sizes 16 and 18. The two images of patterns represent PE outliers at families 17 and 18.

When those patterns with a large selection bias from Figure 9 are plotted showing their bond type, no correlation exists between bond type and family size as seen in Figure 17.

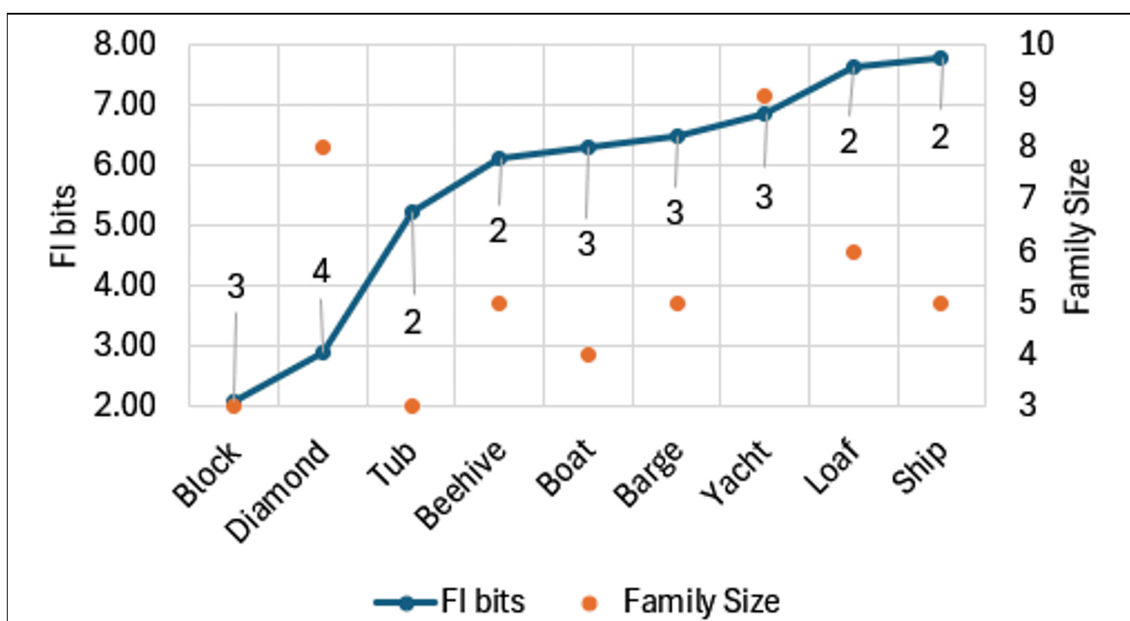


Figure 17. Bond Type and Family Size. Each pattern has a data label showing the associated bond type.

Based on the data from Figure 17, neither family size nor bond type have a correlation with the gradual decrease in frequency of occurrence of these patterns and the rise in FI value in bits.

6. Discussion and Outlook

This research's goal was to provide tools to study the complexity of self-organizing systems. Perennial evolution, as an example of such a system, drives a cellular automaton to longer phases of self-organization due to the catalytic and attractive effect of a permanently active cell interacting with canonical cells governed by the STR algorithm. The resulting creation in this CA of 351 unique geometric patterns – two-dimensional analogues of molecules – demonstrated many examples of self-organized behavior. But did the model, as examined by FI theory, shed any light on that hard-to-define term complexity? I believe it did by restricting the meaning of complexity to the function performed by the automaton, i.e., the creation of stable patterns defined by a minimum size and the requirement that components be interconnected. FI theory, when applied to PE, provided a definition of complexity based on two key metrics:

1. Frequency of Occurrence (FIDR): this metric established a range for the probability of occurrence by function of each pattern. That function was the creation of stable still life patterns with ≥ 3 canonical cells bound to the single perennial cell from the total of 401,275 random soups of which 229,570 created a pattern.

2. Information in Bits (FI): this metric converted the FIDR to an information-based value that allows a numerical comparison of information content for patterns generated by the algorithmic STR.

It was determined, based on the number of copies of a pattern, that these two metrics are correlated with the typical number of generations of evolution that constructed a pattern based on the Construction Index (CI). Limits to this correlation were also described: as the number of copies of a pattern decreases, the determination of a typical value for construction generations is problematic.

The relationship of CI to the Assembly Index (AI) of the theory invented by astrobiologists to aid in discriminating biosignatures in the complex signals from exoplanets deserves a comment since this author has been previously involved in astrobiological research^[30] and the name “Construction Index” pays homage to that work. Assembly Theory, with its origin in astrobiology^{[31][32][33][34]}, has been recently modified to apply to cellular automata^[35] and is called AssemblyCA. In the application of this theory to CA, objects – what in this research are called patterns – are built from preexisting objects and the assembly of these objects into more complex objects is directed by an agent. The AI 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 complex systems analysis.

In work not yet reported, I created an additional 100 subsets of data that increased the total initial configurations by 96,543. The result of this increase did not change the relative occurrence frequency of the reported patterns described in this work but did create 35 new patterns. Of these 35 new discoveries 18 were very similar to what had been previously created with a single PE bond; some were simply new isometric forms of previous patterns. In the work reported here no patterns were detected in family 16 but with additional datasets, two patterns were created for this family. Thus, the capacity for diversity in perennial evolution appears limited, even for those patterns with only a single bond creating stability. While the FI values of patterns will increase with additional subsets of initial configurations, it is not certain that their diversity increases despite an increase in complexity as quantified by Functional Information theory.

An example of future ideas that could be explored with the techniques developed in this research, consider two patterns as analogues of primordial minerals that bond with each other. This would be analogous to the way minerals become more complex^[24] over long periods of geological time during which thermodynamic and gravimetric churning introduce phase changes^[36]. These phase changes could be represented as the conversion of primordial mineral patterns to a fully perennial form as shown in Figure 18 and allow the patterns to interact *without* any soup of canonical cells, forming new stable structures.

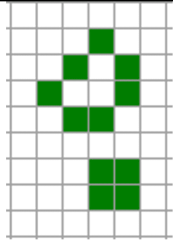
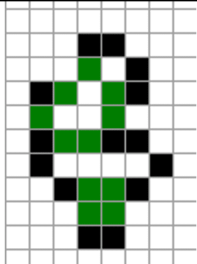
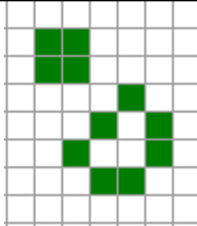
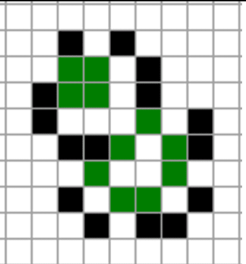
(a) Initial State	(b) End State after 3,099 Generations	(c) Initial State	(d) End State after 78,559 Generations
			

Figure 18. A Loaf and Block as Interacting Patterns. The relative location of the two patterns shown in (a) vs (c) results in different end state structures as well as differences in the length of self-organized construction as measured in generations. Note that the placement of these patterns near each other results in triggering the STR that generates canonical cells which initiate evolution.

This project's model has demonstrated one technique of using cellular automata to quantify self-organized complexity, and future possibilities remain to be explored. Good hunting to all those who share in this quest.

Appendix

A1. Dynamics of Construction

An example of the attractive effect of a perennial cell is demonstrated during the construction of the Block, the most frequently occurring pattern with 94,219 copies. 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 percentage of evolutionary time a cell is active illustrates the attraction as shown in Figure A1-1.

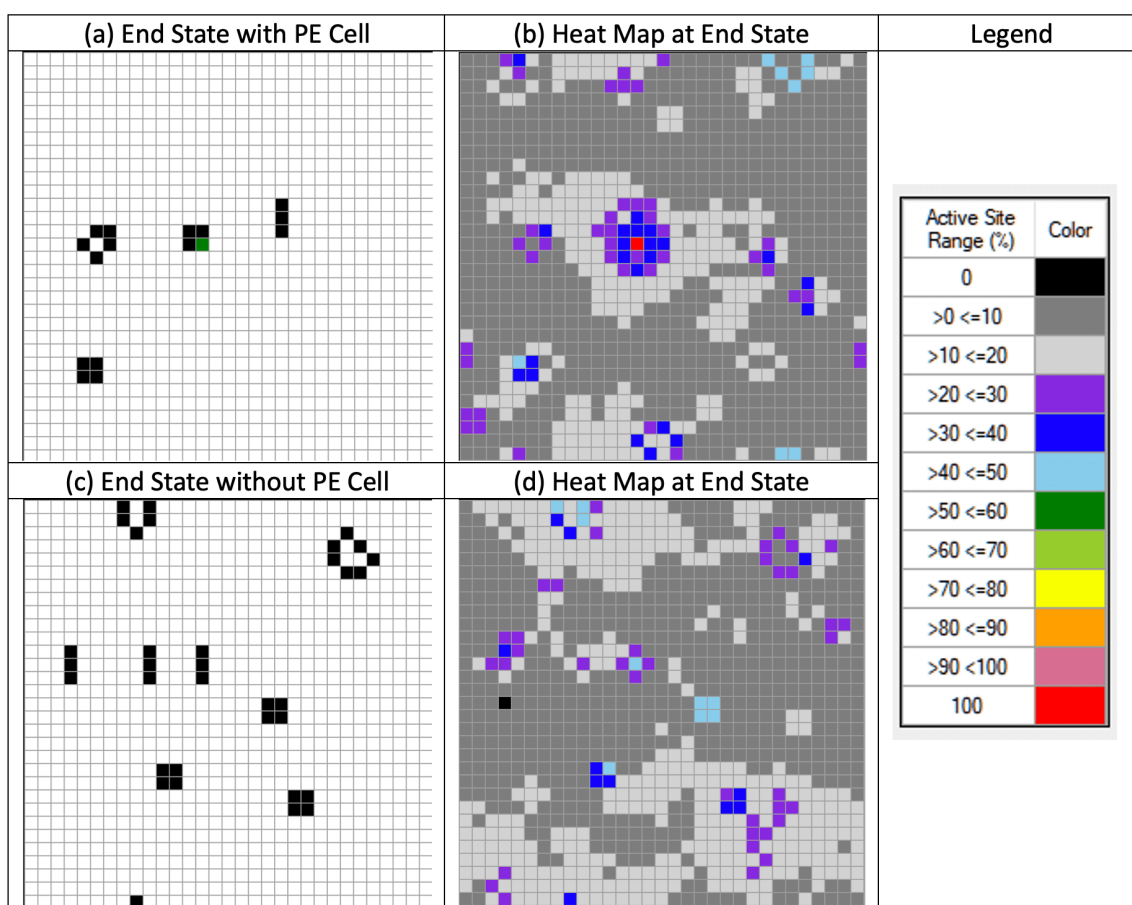
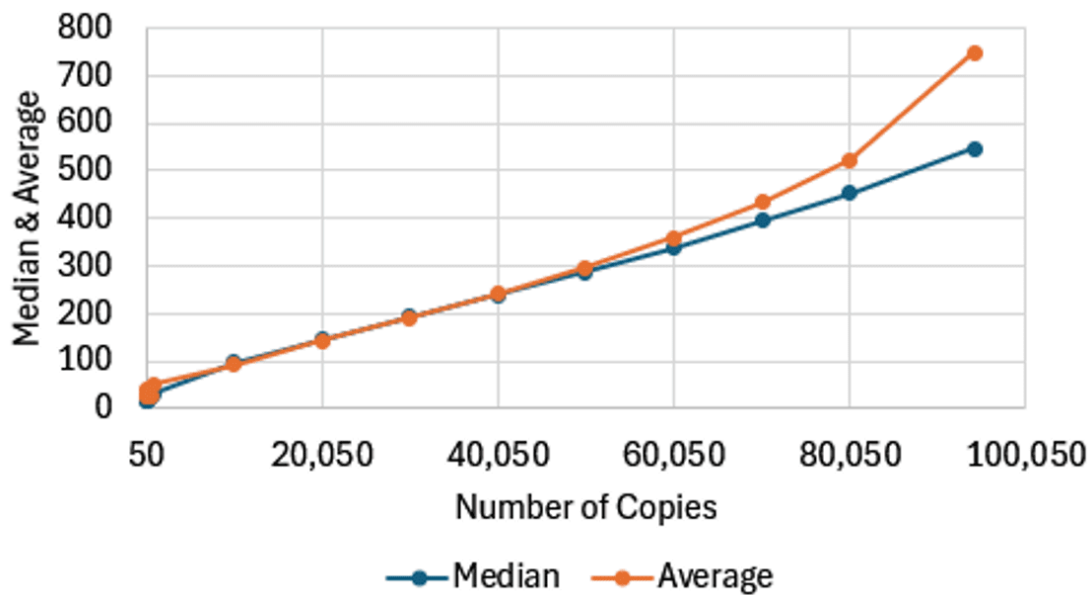


Figure A1.1. Heat Maps. The end state of evolution with and without the PE cell is shown respectively in (a) and (c). Both soups reached an end state at 600 generations. In (a) the Block formed at generation 542. In (b) the region around the PE Cell, located at grid cell number 447 and red in color, is active during all generations and has a much higher number of neighboring cells active than around the same location in (d).

The high activity around the PE cell is due to its attractive effect despite the average overall percentage of all cell activity of both soups being nearly equal: in Figure A1.1 (a) it was 9.6% and in (c) it was 9.44%.

A2. Median CS vs Average CS

For the distribution of CS values shown for the Block in Figure 12 of the manuscript, the median and average values vary as a function of the number of copies. For all 94,219 copies the final median value is 546 and the final average value is 749 generations. The absolute median and average CS values are shown in the graph below as they approach their respective final values.



In this graph, as the number of copies increases toward the final number (94,219), extreme values in the tail of the distribution pull the average higher.

The conclusion from this examination of the Block's relationship between the median and average values is that the median is a more appropriate typical value for the CS value from which a CI value is calculated. Because the CS distribution for other patterns is like the Block's, then the median will also be the metric value of choice.

Regarding the small CS variability of 550 ± 12 among the patterns with $\geq 1,000$ copies (Figure 9 & Table 1), a question arose that a constant value might suffice for these patterns, or that if more configurations were executed, so that the number of copies increased, might the median CS values converge? To answer this question, additional configurations were run for the Block, Diamond and Tub – the patterns with the largest frequency of occurrence. The number suffix in the pattern names in the table below have the following meaning: “1” is for this current research with 410,275 total random soups and “2” is for a total of 512,324. The three patterns had an increase in their number of copies as shown in the table.

	Block 1	Block 2	Diamond 1	Diamond 2	Tub 1	Tub 2
Copies	94,219	120,125	53,337	68,226	10,696	13,659
Median CS	546	548	560	558	543	541

The values have not converged: this also establishes the fact that each unique pattern should be represented by

its own CS value based on its total number of copies, and not a constant median CS for all patterns with $\geq 1,000$ copies.

Supplementary data

Available [here](#).

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 AI in Research

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.

Acknowledgements

The author thanks Qeios reviewers Dr. Al-Dabbagh, Serge Dolgikh, Fabrizio Browen, and Szymon Lukaszuk for helpful 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](https://doi.org/10.1016/j.physrep.2024.04.001).
2. [△]Deng X, Shao Y, Song J, Wu H (2022). "Traffic Flow Simulation of Modified Cellular Automata Model Based on Producer-Consumer Algorithm." *PeerJ Comput Sci.* **8**:e1102. doi:[10.7717/peerj-cs.1102](https://doi.org/10.7717/peerj-cs.1102). S2CID [252417533](https://pubmed.ncbi.nlm.nih.gov/252417533/).

3. ^ΔFrisch U, Hasslacher B, Pomeau Y (1986). "Lattice-Gas Automata for the Navier-Stokes Equation." *Phys Rev Lett.* **56**(14):1505. <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.56.1505>.
4. ^ΔBatty M (2007). *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*. The MIT press. <https://dl.acm.org/doi/abs/10.5555/1543541>.
5. ^ΔWhite SH, Del Rey AM, Sánchez GR (2007). "Modeling Epidemics Using Cellular Automata." *Appl Math Comput.* **186**(1):193–202. doi:[10.1016/j.amc.2006.06.126](https://doi.org/10.1016/j.amc.2006.06.126).
6. ^ΔLangton CG (1984). "Self-Reproduction in Cellular Automata." *Physica D.* **10**(1-2):135–144. doi:[10.1016/0167-2789\(84\)90256-2](https://doi.org/10.1016/0167-2789(84)90256-2).
7. ^ΔHutton TJ (2002). "Evolvable Self-Replicating Molecules in an Artificial Chemistry." *Artif Life.* **8**(4):341–356. doi:[10.1162/106454602321202417](https://doi.org/10.1162/106454602321202417).
8. ^ΔFlitney AP, Abbott D (2010). "Towards a Quantum Game of Life." In *Game of Life Cellular Automata*. London: Springer London. pp. 465–486. https://www.researchgate.net/profile/Adrian-Flitney/publication/236109736_Towards_a_Quantum_Game_of_Life/links/55666d3208aec22682ff1b0f/Towards-a-Quantum-Game-of-Life.pdf.
9. ^ΔTurney PD (2021). "Evolution of Autopoiesis and Multicellularity in the Game of Life." *Artif Life.* **27**(1):26–43. doi:[10.1162/artl.000334](https://doi.org/10.1162/artl.000334).
10. ^ΔCurry S (2015). "Structural Biology: A Century-Long Journey into an Unseen World." *Interdiscip Sci Rev.* **40**(3):308–328. doi:[10.1179/0308018815Z.0000000000120](https://doi.org/10.1179/0308018815Z.0000000000120).
11. ^Δ^bRainwater JH (2024). "Self-Organization and Phase Transitions in Driven Cellular Automata." *Artif Life.* **30**(3):302–322. doi:[10.1162/artl.000437](https://doi.org/10.1162/artl.000437).
12. ^ΔGardner M (1970). "Mathematical Games." *Sci Am.* **223**(4):120–123. doi:[10.1038/scientificamerican1070-120](https://doi.org/10.1038/scientificamerican1070-120).
13. ^ΔHazen RM, Griffin PL, Carothers JM, Szostak JW (2007). "Functional Information and the Emergence of Biocomplexity." *Proc Natl Acad Sci U S A.* **104**(suppl_1):8574–8581. doi:[10.1073/pnas.0701744104](https://doi.org/10.1073/pnas.0701744104).
14. ^ΔBeer RD (2020). "An Investigation into the Origin of Autopoiesis." *Artif Life.* **26**(1):5–22. doi:[10.1162/artl.000307](https://doi.org/10.1162/artl.000307).
15. ^Δ^bSzostak JW (2003). "Functional Information: Molecular Messages." *Nature.* **423**(6941):689–689. doi:[10.1038/423689a](https://doi.org/10.1038/423689a).
16. ^ΔJohnson 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](https://doi.org/10.1098/rsta.2019.0301).
17. ^ΔHazen R, Morison S (2020). "An Evolutionary System of Mineralogy. Part I: Stellar Mineralogy (>13 to 4.6 Ga)." *Am Mineral.* **105**(5):627–651. doi:[10.2138/am-2020-7173](https://doi.org/10.2138/am-2020-7173).
18. ^ΔSmith E, Morowitz HJ (2016). *The Origin and Nature of Life on Earth: The Emergence of the Fourth Geosphere*. Cambridge: Cambridge University Press.

19. ^ΔBak P, Chen K, Creutz M (1989). "Self-Organized Criticality in the Game of Life." *Nature*. 342(6251):780–782. <http://www.nature.com/articles/342780a0.pdf>.
20. ^ΔReia SM, Kinouchi O (2014). "Conway's Game of Life Is a Near-Critical Metastable State in the Multiverse of Cellular Automata." *Phys Rev E*. 89(5):052123. doi:10.1103/PhysRevE.89.052123.
21. ^ΔGarcia JBC, Gomes MAF, Jyh TI, Ren TI, Sales TRM (1993). "Nonlinear Dynamics of the Cellular-Automaton "Game of Life"." *Phys Rev E*. 48(5):3345. doi:10.1103/PhysRevE.48.3345.
22. ^ΔGould SJ (1989). *Wonderful Life: The Burgess Shale and the Nature of History*. New York, NY: W. W. Norton & Co.
23. ^ΔHazen 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.
24. ^Δ^bHazen RM, Wong ML (2024). "Open-Ended Versus Bounded Evolution: Mineral Evolution as a Case Study." *Pnas Nexus*. 3(7):page 248. doi:10.1093/pnasnexus/pgae248.
25. ^ΔWong ML, Cleland CE, Arend D Jr, Bartlett S, Cleaves HJ, Demarest H, Prabhu A, Lunine JJ, Hazen RM (2023). "On the Roles of Function and Selection in Evolving Systems." *Proc Natl Acad Sci U S A*. 120(43):e2310223120. doi:10.1073/pnas.2310223120.
26. ^ΔAdami C (2016). "What Is Information?" *Philos Trans R Soc A*. 374(2063):20150230. doi:10.1098/rsta.2015.0230.
27. ^ΔAdami C, Cerf NJ (2000). "Physical Complexity of Symbolic Sequences." *Physica D*. 137(1–2):62–69. <https://www.sciencedirect.com/science/article/abs/pii/S0167278999001797?via%3Dihub>.
28. ^ΔBhatti SH, Hussain S, Ahmad T, Aslam M, Aftab M, Raza MA (2018). "Efficient Estimation of Pareto Model: Some Modified Percentile Estimators." *PLoS One*. 13(5):e0196456. doi:10.1371/journal.pone.0196456.
29. ^ΔSornette D (2009). "Dragon-Kings, Black Swans and the Prediction of Crises." *arXiv preprint*. doi:10.48550/arXiv.0907.4290.
30. ^ΔHoehler 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, Lafuent e 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*. doi:10.1089/ast.2024.0107.
31. ^ΔCronin L, Walker SI (2016). "Beyond Prebiotic Chemistry." *Science*. 352(6290):1174–1175. doi:10.1126/science.aaf6310.
32. ^ΔMurray A, Marshall S, Cronin L (2018). "Defining Pathway Assembly and Exploring Its Applications." *arXiv preprint arXiv:1804.06972*. doi:10.48550/arXiv.1804.06972.

33. [^]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. doi:[10.1038/s41586-023-06600-9](https://doi.org/10.1038/s41586-023-06600-9).
34. [^]Gonçalves D (2024). "Rethinking Life and Predicting Its Origin." *Theory Biosci*. 143(3):205–215. doi:[10.1007/s12064-024-00420-9](https://doi.org/10.1007/s12064-024-00420-9).
35. [^]Patarroyo KY, Sharma A, Walker S, Cronin L (2023). "AssemblyCA: A Benchmark of Open-Endedness for Discrete Cellular Automata." In *Second Agent Learning in Open-Endedness Workshop*. doi:<https://openreview.net/forum?id=5cEQ4ZOsIN>.
36. [^]Schwartzman DW, Shore SN, Volk T, McMenamin M (1994). "Self-Organization of the Earth's Biosphere-Geochemical or Geophysiological?" *Origins Life Evol Biosphere*. 24:435–450. doi:[10.1007/BF01582019](https://doi.org/10.1007/BF01582019).

Supplementary data: available at <https://doi.org/10.32388/HX4713.4>

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.