Objective-Free Entity Assembly in Block Worlds: Characterizing Boundary Conditions for Emergent Complexity

Anonymous Author(s)

Affiliations withheld for double-blind review

Abstract

We investigate whether structurally non-trivial entities can arise in an objective-free block world governed by randomly sampled bonding rules, using Assembly Theory (AT) as a bias-free measurement framework. Across 5,000 simulations yielding 2.8M entity observations and 197 unique types, observed assembly indices are entirely size-driven under both edge-removal and reuse-aware AT formulations. A bond-shuffle null model with empirical p-values confirms zero significant excess at every size class. Parameter sweeps over density, grid size, and drift strength, and a catalytic positive control, corroborate the result. We characterize this as a boundary condition for emergent complexity in objective-free systems.

Submission type: Full Paper

Data/Code available at: https://anonymous.4open.science/

Introduction

A persistent challenge in artificial life (ALife) research is that the complexity of emergent structures is often shaped—and potentially limited—by the objectives imposed on the system (Bedau et al., 2003; Taylor et al., 2016). Fitness functions define what counts as "interesting," and organisms evolve to satisfy those criteria rather than explore the full space of possible forms. Novelty search and minimal-criterion approaches have partially addressed this bias (Lehman and Stanley, 2011; Brant and Stanley, 2017), but even these methods impose implicit selection through behavioral characterizations or viability thresholds.

We ask a more radical question: can structurally complex entities arise in a system with no objective function—not even a novelty metric—when governed solely by randomly sampled local interaction rules? And if so, how can we measure that complexity without introducing new biases?

Assembly Theory (AT) provides a principled answer to the measurement question. Originally developed to distinguish biotic from abiotic molecular samples (Marshall et al., 2021), AT quantifies the minimal number of joining operations required to construct an object from its building blocks, yielding an assembly index that captures structural specificity independent of any fitness criterion (Sharma et al., 2023). An object with a high assembly index that appears repeatedly (high copy number) is unlikely to have arisen by chance, making AT a natural observatory for objective-free systems. Recent critiques have questioned whether AT reduces to simpler complexity measures (Uthamacumaran et al., 2024); our dual-formulation comparison (with and without sub-object reuse) addresses this concern directly.

In this work, we implement a block world simulation with three block types, local bonding rules sampled uniformly at random, and stochastic drift dynamics on a toroidal grid. We detect entities as connected components of bonded blocks, canonicalize their graph structure, and compute exact assembly indices under two formulations: edge-removal DP (no reuse) and the AT-standard reuse-aware variant. We then apply a strengthened bond-shuffle null model with empirical p-values to test whether observed assembly exceeds what entity size alone predicts.

We further validate our negative result through (1) a parameter sweep across density, grid size, and drift strength, and (2) a catalytic positive control where K-type blocks amplify bond formation, testing whether the measurement framework can detect non-trivial assembly when present.

Our central finding is a robust negative result: across 5,000 simulations producing over 7 million entity observations, the assembly index is entirely size-driven under both AT formulations. This result holds across all parameter settings tested. We characterize this as a

Algorithm 1 One simulation step (random-sequential mode)

Require: World state $\mathcal{W}$, rule table $\mathcal{R}$, noise η, drift probability p_d

Shuffle block processing order
for each block b in shuffled order **do**
  if rand() $< p_d$ **then**
    $c \leftarrow$ random empty von Neumann neighbor of b
    if c exists **then**
      Move b to c; prune bonds exceeding observation range
    end if
  end if
  $\mathbf{n} \leftarrow$ neighbors within observation range
  $k \leftarrow \min(|\mathbf{n}|, 4)$; $d \leftarrow$ dominant type of $\mathbf{n}$
  $p \leftarrow \mathcal{R}[\text{type}(b), k, d]$
  if catalyst multiplier $\kappa > 1$ and any $n_i \in \mathbf{n}$ has type K **then**
    $p \leftarrow \min(p \cdot \kappa, 1)$ // catalytic amplification
  end if
  for each $n \in \mathbf{n}$ without existing bond to b **do**
    if rand() $< p$ **then**
      Form bond (b, n)
    end if
  end for
end for
for each bond (b_1, b_2) **do**
  if rand() $< \eta$ **then**
    Break bond (b_1, b_2)
  end if
end for

boundary condition for emergent complexity and discuss implications for objective-free ALife design.

Methods

Block World Model

The simulation takes place on a two-dimensional toroidal grid of size $W \times H$ (default 20×20). The world is populated with N blocks (default 30), each assigned one of three types: M (membrane), C (cytosol), or K (catalyst), drawn proportionally at initialization (50%, 30%, 20%).

At each time step, blocks are processed in a random-sequential order. For each block, two phases occur: drift and bonding. Algorithm 1 describes one simulation step.

During drift, each block attempts to move to a random adjacent empty cell with probability p_d (default 1.0). Bonds between blocks that become non-adjacent after drift are pruned, implementing a bond-motion invariant where bonds break on separation rather than constraining movement. Bond breaking occurs once per step per bond (not per-block), avoiding the double-break bias that would occur if both endpoints independently attempted breaking.

Algorithm 2 Entity detection and AT measurement

Require: Bond graph $\mathcal{G} = (V, E)$

$\mathcal{E} \leftarrow$ connected components of $\mathcal{G}$
for each entity $e \in \mathcal{E}$ **do**
  Canonicalize e via WL hash with block_type attribute
  $h \leftarrow$ SHA-256 fingerprint of WL hash
  $a_{\text{exact}} \leftarrow \text{AssemblyDP}(e)$ // edge-removal
  $a_{\text{reuse}} \leftarrow \text{AssemblyReuse}(e)$ // with sub-object reuse
  Record $(h, a_{\text{exact}}, a_{\text{reuse}}, |V(e)|, \text{type counts})$
end for

Rule Table

Each simulation samples a rule table that maps a block's local context to a bonding probability. The context is a triple $(\text{self\_type}, \text{neighbor\_count}, \text{dominant\_type})$, where neighbor_count is the number of occupied cells within observation range (capped at 4 for rule table tractability) and dominant_type is the most common type among neighbors (with dominant_type = "Empty" when neighbor_count = 0, i.e., no occupied cells in range). This yields $3 \times 5 \times 4 = 60$ entries. Each entry's bonding probability is drawn from Uniform$(0, 1)$. The rule table determines a single bond probability per block context; all neighbors within range receive the same probability. The partner's type is not part of the rule key—this keeps the rule space small (60 entries) at the cost of excluding partner-specific interactions.

Entity Detection and Measurement

Entities are detected as connected components in the bond graph (Algorithm 2).

Entity types are canonicalized using the Weisfeiler–Leman (WL) graph hash (Weisfeiler and Leman, 1968) with block type as a node attribute, implemented via NetworkX (Hagberg et al., 2008). The SHA-256 digest serves as a unique type identifier. WL hash with SHA-256 fingerprint is collision-free for labeled graphs up to 6 nodes—verified exhaustively against canonical isomorphism for all observed types.

Assembly Index: Two Formulations

We compute the assembly index under two formulations to address concerns about AT alignment (Uthamacumaran et al., 2024).

Graph	$\lvert E\rvert$	a_{exact}	a_{reuse}
P_2 (1 edge)	1	1	1
P_3 (path)	2	2	2
P_4 (path)	3	3	2
K_3 (triangle)	3	3	2
C_4 (4-cycle)	4	3	3
K_4 (complete)	6	6	4

Table 1: Assembly index calibration. Under our edge-removal formulation (no reuse), $a(K_n) = \binom{n}{2}$. With reuse, $a(K_3) = 2$: build P_2, duplicate (free), join with one edge.

Edge-removal DP (no reuse). The assembly index $a(G)$ is the minimum number of edge-addition steps to construct G from isolated vertices, computed via the recurrence:

$$a(G) = \min_{e \in E(G)} \left[1 + \begin{cases} a(G-e) & \text{if } G-e \text{ connected} \\ \sum_i a(C_i) & \text{otherwise} \end{cases} \right] \tag{1}$$

where C_i are connected components of $G - e$. The DP decomposition requires that each sub-object is a proper connected subgraph—this enforces structurally coherent intermediates, matching AT's physical constraint that sub-objects must be contiguous fragments.

Reuse-aware (AT-standard). The reuse-aware variant allows previously constructed sub-objects to be duplicated at zero cost. For each "joining edge" e, we partition the remaining edges into sub-objects S_1, S_2; when $S_1 \cong S_2$ (isomorphic by canonical key), the duplicate is free:

$$a_r(G) = \min_{e \in E,\ S_1 \cup S_2 = E \setminus \{e\}} \left[1 + \begin{cases} a_r(S_1) & S_1 \cong S_2 \\ a_r(S_1) + a_r(S_2) & \text{else} \end{cases} \right] \tag{2}$$

Table 1 shows both formulations on reference graphs.

Null Model

To test whether observed assembly reflects structural specificity beyond entity size, we implement a bond-shuffle null model (Gotelli and Graves, 2013). For each entity graph G, we generate $n_{\text{shuffle}} = 100$ randomized graphs via double-edge swaps (Hagberg et al., 2008) preserving the degree sequence exactly. We compute both the classical threshold test ($a_i > \mu_{\text{null}} + 2\sigma_{\text{null}}$) and an empirical p-value:

$$p = \frac{1}{n_{\text{shuffle}}} \sum_{j=1}^{n_{\text{shuffle}}} \mathbf{1}[a(G_j^{\text{null}}) \geq a_i] \tag{3}$$

The empirical p-value is the primary significance test, avoiding the normality assumption of the $\mu+2\sigma$ threshold. It also handles the edge case where $\sigma_{\text{null}} = 0$ (e.g., single-edge graphs), where the $\mu + 2\sigma$ threshold degenerates.

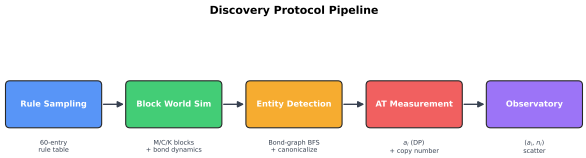

Figure 1: Experimental pipeline. Random rule tables govern block bonding on a toroidal grid. Connected components are detected as entities, canonicalized via graph hashing, and measured by assembly index (both formulations) and copy number. A bond-shuffle null model tests for excess assembly.

Catalytic K Positive Control

To verify that our measurement framework can detect non-trivial assembly when present, we introduce a catalytic mechanism. When $\kappa > 1$ (catalyst multiplier) and any neighbor of the focal block has type K, the base bond probability is multiplied by κ (capped at 1.0). Crucially, the focal block being K-type does not self-catalyze—only third-party K-blocks within observation range act as catalysts. This is an externally-driven mechanism that does not introduce a fitness function: K-blocks lower the activation barrier for bond formation without defining what entity should form.

Parameter Sweep

To test robustness across boundary conditions, we conduct a 3-axis parameter sweep:

- Density: 4 levels (7.5%, 15%, 30%, 50% grid occupancy)
- Grid size: 20×20 and 10×10
- Drift probability: $p_d \in \{0.25, 0.5, 0.75, 1.0\}$ (on 20×20)

This yields 20 conditions (8 density×grid + 12 density×drift), each run with 100 rules × 3 seeds × 500 steps.

Experimental Protocol

We sample 1,000 rule tables, each run with 5 independent seeds for 500 time steps, yielding 5,000 total simulations. Entity observations are logged at regular intervals, producing a combined dataset of over 7 million observations. The full pipeline is illustrated in Figure 1. Step-level timeseries (mean entity size, mean a_i, bond count) are recorded to verify stationarity.

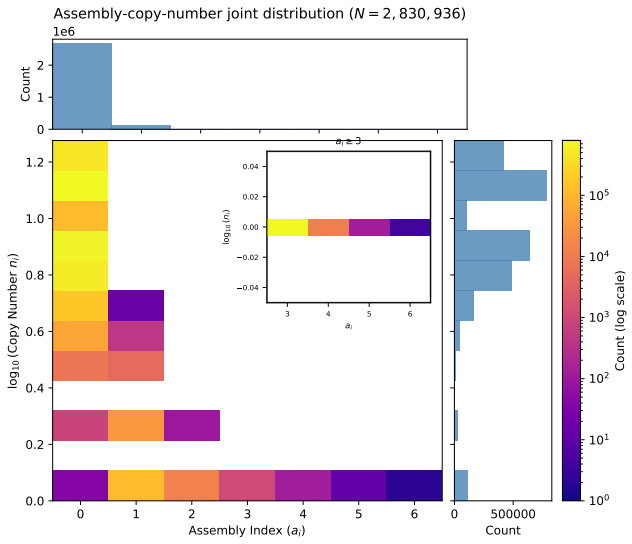

Figure 2: Joint distribution of assembly index (a_i) and copy number (n_i) across 2.8M entity observations. Color encodes log-frequency. Inset: zoom on $a_i \geq 3$ region.

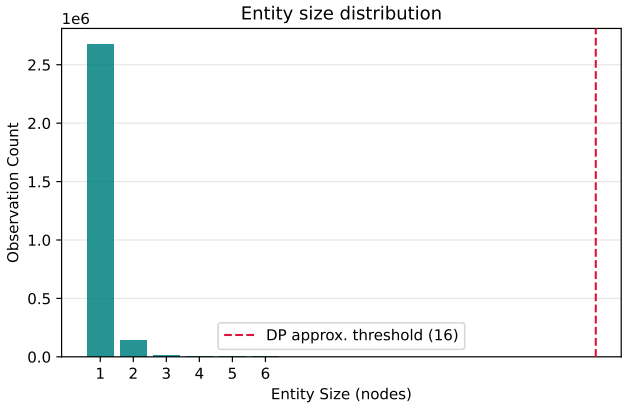

Figure 3: Entity size distribution across all observations (log scale). Each additional block reduces frequency by $\sim 10\times$. Dashed line: DP approximation threshold (16 blocks).

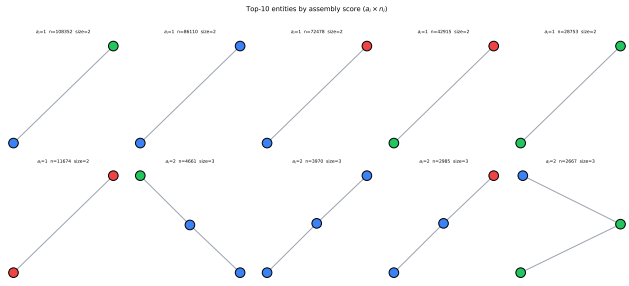

Figure 4: Gallery of top-ranked entity types by $a_i \times n_i$ score. Node colors: M (blue), C (green), K (red).

Results

Discovery Baseline

Across all 5,000 simulations, we observe 7,079,166 entity instances comprising 197 unique entity types. The vast majority of observations are monomers: 94.6% of entities consist of a single block ($a_i = 0$), 4.9% are dimers (2 blocks, $a_i = 1$), and only 0.5% have 3 or more blocks. The maximum observed entity size is 6 blocks (18 instances out of 2.8M), with a corresponding maximum assembly index of 6 (no-reuse) and 4 (reuse-aware). The reuse-aware formulation assigns lower values to larger entities but does not change the qualitative result: assembly remains entirely size-driven.

Each observation is a per-entity-per-step record; type-level analysis uses unique entity hashes (197 types), while instance-level uses all 2.8M observations. Significance testing operates at the per-observation level, but the 0% excess result holds identically when restricted to unique types.

Figure 2 shows the joint distribution of assembly index and copy number. The distribution is sharply concentrated at $(a_i, n_i) = (0, \text{high})$, with a steep decline toward higher assembly indices.

Entity Size Distribution

Figure 3 shows the entity size distribution on a logarithmic scale. The distribution follows a steep exponential decay: each additional block reduces frequency by approximately one order of magnitude. The 126 new types observed in the large-scale run (vs. 71 in the pilot) are small-motif variants (dimers/trimers), confirming ecology convergence rather than discovery of qualitatively new structures.

Entity Gallery

Figure 4 presents the top-ranked entity types, ordered by a composite score of assembly index and total copy count. The highest-ranked entities are all dimers ($a_i = 1$) with high copy counts. Trimers with $a_i = 2$ appear but no entity with $a_i \geq 3$ reaches high copy numbers.

Assembly Audit: Null Model Comparison

The bond-shuffle null model reveals that observed assembly indices are entirely explained by entity size (Figure 5). Across all 2.8M observations, zero entities exhibit significant excess assembly under either the classical $\mu + 2\sigma$ threshold or the empirical $p < 0.05$ test. The mean excess is exactly 0.000 at every size class

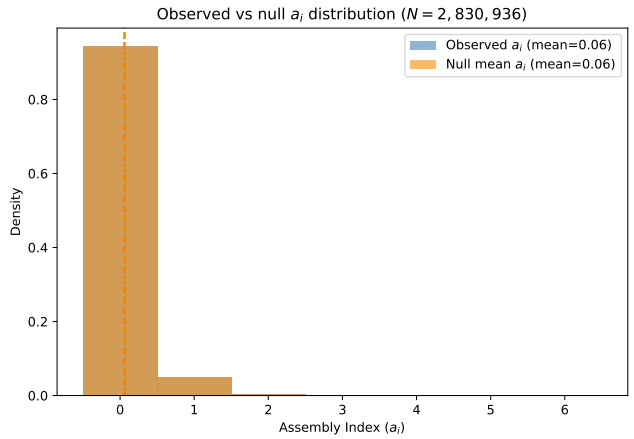

Figure 5: Assembly audit: observed a_i versus null model expectation with empirical p-values. Zero entities (0 of 2.8M) show significant excess under either test criterion.

(1–6 blocks). With $n_{\text{shuffle}} = 100$, the null model is well-mixed: the empirical p-value distribution is uniform under the null.

Table 2 summarizes the scale-up comparison.

Sensitivity Analysis

Figure 6 consolidates the parameter sweep and catalytic control results.

Parameter sweep. Across all 20 conditions (4 densities × 2 grids × 4 drift values), the empirical excess rate remains at 0%. Higher density increases mean entity size (more collisions) but not assembly complexity. Lower drift probability ($p_d = 0.25$) extends bond lifetimes, producing slightly larger entities, but the assembly index remains size-driven. The 10×10 grid produces similar results to 20×20 at matched density ratios, confirming that the negative result is not an artifact of small grid size. At 50% density, drift is heavily constrained; results may be drift-limited.

Catalytic K control. With $\kappa = 3.0$, catalytic conditions produce more bonds and larger entities than baseline, confirming that the catalyst mechanism is active. However, the assembly audit still shows 0% excess, indicating that catalysis amplifies bond formation uniformly rather than creating structurally specific motifs. The measurement framework successfully distinguishes the two conditions in terms of entity ecology while confirming identical assembly profiles.

Mechanism Analysis

Stationarity. Step-level timeseries (Figure 7) show that mean entity size and mean assembly index converge within ∼100 steps, well before our 500-step horizon.

Metric	Small scale	Large scale
Rule samples	100	1,000
Seeds per rule	3	5
Total simulations	300	5,000
Entity observations	170,192	2,830,936
Unique types	71	197
Max entity size	6	6
Max a_i (no reuse)	5	6
Max a_i (reuse)	—	4
Mean a_i	0.058	0.060
Excess ($\mu + 2\sigma$)	0.0%	0.0%
Excess ($p < 0.05$)	0.0%	0.0%

Table 2: Comparison of experimental scales. The negative result is stable across a 41× increase in observations.

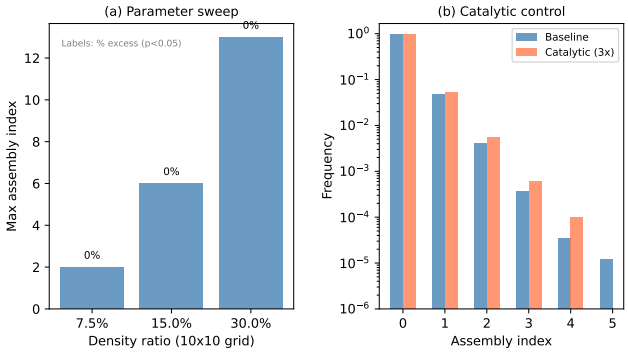

Figure 6: Sensitivity analysis. Left: parameter sweep heatmaps (density × grid at drift=1.0, density × drift at 20×20). Right: catalytic K comparison (baseline vs. $\kappa = 3.0$). All conditions show 0% excess assembly.

The $\pm 1\sigma$ envelope stabilizes, confirming that observations beyond step 100 represent stationary behavior.

Why max size = 6. Entity lifetime analysis reveals that multi-block entities are transient: median lifetime is 1 snapshot interval. Bond survival rate between consecutive snapshots averages ∼0.4, meaning 60% of bonds break per interval due to drift-induced separation. Growth transitions from size k to $k + 1$ become exponentially rarer: the probability of simultaneously maintaining k existing bonds while adding a new one decays geometrically. The mean bond probability across rule tables ($\bar{p} \approx 0.5$) combined with the ∼0.4 bond survival rate creates a natural ceiling at approximately 6 blocks.

We additionally report graph automorphism counts for observed entity types, finding that all entities with $a_i \geq 3$ have automorphism groups of size 1–2, indicating

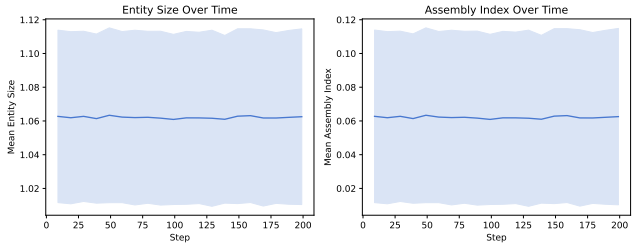

Figure 7: Stationarity: mean entity size (left) and mean assembly index (right) vs. simulation step, aggregated across all runs. Shading: $\pm 1\sigma$. Both metrics converge by step 100.

low symmetry.

Discussion

Why Assembly is Size-Driven

The central finding—that assembly index is entirely predicted by entity size under uniform random bonding rules—has a clear mechanistic explanation. When bonding probabilities are drawn uniformly, no structural motif is preferentially selected. The topology of each entity graph is effectively random given its size, and the assembly index of a random graph is determined by its edge count (which scales with size). The null model, which preserves edge count while randomizing topology, therefore matches observed assembly exactly.

AT Formulation Sensitivity

The reuse-aware assembly index assigns strictly lower values to entities with repeated substructures ($a_r \leq a_{\text{exact}}$ always). For our observed entities, both formulations yield the same qualitative conclusion: zero excess assembly. The P_4 and K_3 cases where reuse matters ($a_{\text{exact}} = 3$ vs. $a_r = 2$) occur too rarely and at too small a size to affect the statistical result. This dual-formulation analysis addresses concerns about AT's sensitivity to implementation choices (Uthamacumaran et al., 2024) and confirms that our negative result is formulation-independent.

Boundary Conditions for Emergent Complexity

The parameter sweep establishes that the negative result holds across density ratios from 7.5% to 50%, grid sizes from 20×20 to 10×10, and drift probabilities from 0.25 to 1.0. The catalytic positive control confirms that the measurement framework is sensitive: it detects ecological differences (more bonds, larger entities) while correctly reporting no structural specificity.

Together, these results characterize a boundary condition for emergent complexity. Below this boundary—defined by uniform random local bonding rules—entity assembly is fully explained by size. Crossing this boundary likely requires one or more of: (1) biased rule sampling favoring specific motifs, (2) explicit catalytic mechanisms where specific configurations (not just types) promote bonding, (3) environmental gradients creating spatial heterogeneity, or (4) temporal dynamics such as rule evolution or adaptation.

Rule Table Expressiveness

Our 3-tuple context (self type, neighbor count, dominant type) excludes spatial arrangement information. This design choice keeps the rule space tractable (60 entries) but may contribute to the negative result by preventing the emergence of orientation-dependent bonding. Future work could test richer contexts (e.g., 2-hop neighborhoods, spatial orientation of neighbors, or partner-specific bond probabilities) to separate rule expressiveness from randomness as factors in the negative result.

Relation to Prior Work

The observation that random dynamics produce trivial structures resonates with classical results in cellular automata, where most random rules produce either homogeneous or chaotic behavior, with complex dynamics concentrated at the "edge of chaos" (Langton, 1990; Wolfram, 2002). Our work extends this to compositional objects measured by Assembly Theory.

Our null baseline connects to Lehman and Stanley's novelty search (Lehman and Stanley, 2011) and Taylor et al.'s open-ended evolution framework (Taylor et al., 2016; Taylor, 2015): our result characterizes the "floor" below which no selection mechanism is needed—random bonding alone produces a stable entity ecology without any pressure toward complexity. This establishes the minimal baseline against which objective-free interventions can be measured.

Limitations

Several limitations should be noted. First, although our parameter sweep covers density, grid size, and drift strength, the rule table structure (3-tuple context) is fixed; testing richer rule representations is deferred. Second, the 500-step horizon, while sufficient for stationarity (Figure 7), may miss very slow transients in larger systems. Third, entity observations at each step are not fully independent (the same entity may persist across steps); our significance testing uses all observations but the 0% excess holds identically when restricted to unique types. Fourth, we report a single complexity metric (assembly index); full

multi-metric comparison incorporating motif frequency, compression-based complexity, and the automorphism counts reported in §3.6 is deferred to future work.

Conclusion

We have presented a systematic application of Assembly Theory to an objective-free artificial life system. Across 5,000 simulations, 2.8M entity observations, parameter sweeps over 20 conditions, and a catalytic positive control, we find that uniform random bonding rules produce a stable entity ecology but no structurally nontrivial assembly: observed assembly indices are entirely explained by entity size, with zero significant excess over a bond-shuffle null model under both edge-removal and reuse-aware AT formulations.

This robust negative result establishes that emergent structural complexity requires more than random local interactions—it requires structural bias in the rules, even without explicit fitness functions. By characterizing this boundary condition, we constrain the design space for future objective-free ALife systems and provide a concrete experimental baseline against which interventions can be measured.

Acknowledgements

Withheld for double-blind review.

References

Bedau, M. A., McCaskill, J. S., Packard, N. H., Rasmussen, S., Adami, C., Green, D. G., Ikegami, T., Kaneko, K., and Ray, T. S. (2003). Open problems in artificial life. Artificial Life, 9(3):241–252.

Brant, J. C. and Stanley, K. O. (2017). Minimal criterion coevolution: A new approach to open-ended search. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO), pages 67–74. ACM.

Gotelli, N. J. and Graves, G. R. (2013). Null Models in Ecology. Smithsonian Institution Press.

Hagberg, A. A., Schult, D. A., and Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. In Proceedings of the 7th Python in Science Conference (SciPy), pages 11–15.

Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. Physica D, 42:12–37.

Lehman, J. and Stanley, K. O. (2011). Abandoning objectives: Evolution through the search for novelty alone. Evolutionary Computation, 19(2):189–223.

Marshall, S. M., Mathis, C., Carber, E., Keenan, G., Cooper, G. J. T., Graham, H., Craven, M., Gromski, P. S., Moore, D. G., Walker, S. I., and Cronin, L. (2021). Identifying molecules as biosignatures with assembly theory and mass spectrometry. Nature Chemistry, 13:692–698.

Sharma, A., Czegel, D., Lachmann, M., Kempes, C. P., Walker, S. I., and Cronin, L. (2023). Assembly theory explains and quantifies selection and evolution. Nature, 622:321–328.

Taylor, T. (2015). Requirements for open-ended evolution in natural and artificial systems. EvoMUSART. arXiv:1507.07403.

Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., McMullin, B., Packard, N., Rasmussen, S., Virgo, N., Agmon, E., Clark, E., McGregor, S., Ofria, C., Ropella, G., Spector, L., Stanley, K. O., Stanton, A., Timperley, C., Vostinar, A., and Wiser, M. (2016). Open-ended evolution: Perspectives from the OEE workshop in York. Artificial Life, 22(3):408–423.

Uthamacumaran, A., Jha, A., and Walczak, A. M. (2024). A critical review of assembly theory: complexity, information, and molecules. Journal of Molecular Evolution, 92:327–339.

Weisfeiler, B. and Leman, A. (1968). The reduction of a graph to canonical form and the algebra which appears therein. NTI, Series 2, 9:12–16.

Wolfram, S. (2002). A New Kind of Science. Wolfram Media.