

Article

Not peer-reviewed version

Why Life Exists

Michael Timothy Bennett *

Posted Date: 3 March 2026

doi: 10.20944/preprints202603.0203.v1

Keywords: origins of life; ool; stack theory; functional information; weakness



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Why Life Exists

Michael Timothy Bennett

Australian National University, Australia; m@michaeltimothybennett.com

Abstract

Functional information measures how rare functional configurations are. Wong and colleagues argue that selection should drive a *law of increasing functional information*. This is often read as a claim that complexity must increase. We give a cleaner interpretation, which is that survivors tend to be the systems that did not overcommit. We model a system as a policy π , meaning a bundle of commitments expressed in a finite embodied vocabulary. New selection pressures arrive as a set of future requirements drawn from the unobserved outcome set U . A currently viable policy leaves an unobserved buffer $B_\pi \subseteq U$ of outcomes it still permits. Under a maximally ignorant novelty model, the survival probability of π is exactly $2^{|B_\pi| - |U|}$. Under any exchangeable novelty prior, survival remains monotone in $|B_\pi|$. So persistence favours weaker constraints on function, where weakness counts the compatible completions left open. We define degree of function as survival probability and functional information as Hazen and Szostak rarity among currently viable policies. Conditioning on persistence reweights the population toward larger buffers, hence higher functional information. This yields a formal version of Wong's law under explicit assumptions. In fully enumerated toy worlds, weakness maximisation improves mean log survival probability by 1.674 bits relative to random choice. Weakness and simplicity are not the same thing. Weakness helps a system persist under novelty, because it keeps more futures compatible. Simplicity can help a system persist because there is less to break. That obviates the need for repair. Complexity requires self-repair to persist, increasing weakness. Life is persistent complexity. In between complex life and simple nonlife is the void of the unviable; complexity which is not alive.

Keywords: origins of life; ool; stack theory; functional information; weakness

Significance Statement

Wong and colleagues proposed a law of increasing functional information in evolving systems. The phrase is easy to misread as a law of increasing complexity. Here we give a mathematical version of the claim that avoids that complexity trap. We model a system as a bundle of commitments expressed in an embodied vocabulary (constraints on function). We prove that if future demands are not known in advance, then selection for persistence biases survivors toward keeping more untested outcomes compatible. That bias increases functional information in the standard Hazen and Szostak sense. The only invariant is not simplicity but weakness, meaning how many further commitments remain available while staying correct.

Main Text

The Law that is Not About Complexity

Wong and colleagues propose that evolving systems tend to increase functional information under selection [1]. A recent exchange frames this as a question about complexity [2,3]. We take a different view. Selection does not need to reward complexity. Selection needs to reward survival, and evolution favours keeping one's options open.

Our result is a survivorship bias theorem. Conditioning a population on persistence reweights it toward policies that stayed compatible with whatever the future demanded. That reweighting

increases expected functional information inside the viable set. Importantly, it does not require any monotone increase in complexity.

Survival is a problem of generalising beyond what you have already been tested on. Generalisation is only possible because the future is not fully specified by the past. That means there is an unobserved region that still contains many possible constraints. A system persists when later constraints land inside what it already left compatible.

The core claim of this paper is that if you do not know what the future will demand, you should not overcommit. Embodiment can be represented as a formal language of constraints [4]. In that formal language the right target for minimising commitment is constraint weakness. Maximising weakness means maximising the size of the completion set left open by your current commitments.

A Minimal Model of Commitments and Novelty

We use a lightweight simplified version of Stack Theory [5], beginning with a finite embodied language L . L is a set of implementable bodily outcomes for body can make by interacting with its environment (called “statements” in related work [5]). Asserting a particular outcome is asserting a constraint on possible worlds. Think of L as a set of partial specifications of possible worlds in which a particular body exists.

Policies and completions.

Asserting an outcome rules some of the other outcomes in and rules some other outcomes out, because it rules in or out possible worlds. For example, if a human raises their arm, then all possible worlds from that point in time on begin with that person’s arm raised. A completion of an outcome is any other outcome that implies the first. Assume a statement π . Its completion set is

$$\text{Ext}(\pi) := \{y \in L \mid \pi \subseteq y\}.$$

Its weakness is

$$w(\pi) := |\text{Ext}(\pi)|.$$

To embody a particular outcome and constrain possible worlds is to adopt a policy. If π represents your present embodied commitments or policy, then $\text{Ext}(\pi)$ is everything compatible with those commitments. Weakness is the number of compatible completions. Higher weakness means fewer promises, or more degrees of freedom remaining. The completion set can also be interpreted as a formal counterpart of Kauffman’s adjacent possible [6], which is the space of configurations reachable from the current state. Weakness measures its size.

Observed and Unobserved Outcomes

Let α be the set of outcomes you have already been tested on (a history of outcomes you have survived). We write the unobserved region as

$$U := L \setminus \alpha.$$

For any policy π , define its unobserved buffer

$$B_\pi := \text{Ext}(\pi) \cap U.$$

Interpretation. U is everything that could matter later but has not mattered yet. B_π is the part of your option volume that lives in the future. This is where surprises arrive.

Novelty as Future Requirements

A novelty set is a subset $S \subseteq U$ of unobserved outcomes that become required later. A novelty prior is a distribution P over subsets $S \subseteq U$. We call P exchangeable when it depends only on $|S|$ and

not on which elements of U are in S .

Interpretation. Exchangeable means you do not know which surprises will matter, only how many of them show up. The position that true novelty is not prestatable aligns with arguments that biological evolution is not governed by entailing laws [7].

Survival Probability Equals Option Volume

Assume π has not yet failed on the observed set α , meaning $\alpha \subseteq \text{Ext}(\pi)$. Under novelty set S , the policy survives if it did not preclude any newly required outcome. In this model that condition is

$$S \subseteq B_\pi.$$

Interpretation. You die when the future demands something you forbade with your embodied policy.

Proposition 1 (Uniform ignorance survival law). *Assume a maximally uninformative novelty prior where every subset $S \subseteq U$ is equally likely. Then the survival probability of a currently correct policy π is*

$$\mathbb{P}(\pi \text{ survives}) = \frac{2^{|B_\pi|}}{2^{|U|}} = 2^{|B_\pi| - |U|}.$$

Interpretation. Under complete ignorance, you survive in proportion to how much untested future you left compatible. Every extra compatible unobserved outcome doubles your survival probability. There is no complexity term in this expression.

Proposition 2 (Exchangeable monotonicity). *For any exchangeable novelty prior P , the survival probability $\mathbb{P}_P(S \subseteq B_\pi)$ is a nondecreasing function of $|B_\pi|$.*

Interpretation. If your only information is how many surprises arrive, then leaving more room for surprises cannot hurt. In set terms, survival is the event $S \subseteq B_\pi$. A larger buffer means more possible novelty sets are contained in it. This is a novelty analogue of survival of the flattest [8], where genotypes on broader fitness plateaus outcompete faster replicators at high mutation rates. Here the “flatness” is the size of the unobserved buffer rather than the density of neutral neighbours in sequence space.

From Option Volume to Functional Information

Hazen and Szostak define functional information as

$$I(F) = -\log_2\left(\frac{M(F)}{N}\right),$$

where N is the total number of configurations and $M(F)$ is the number of configurations with degree of function at least F [9,10].

We apply that definition inside the currently viable set. Fix the observed outcomes α and let Π be the set of policies that remain compatible with α , meaning $\alpha \subseteq \text{Ext}(\pi)$ for every $\pi \in \Pi$. Define the degree of future function of a policy by its survival probability under the novelty prior

$$F(\pi) := \mathbb{P}_P(S \subseteq B_\pi).$$

Define

$$M(\pi) := \{\pi' \in \Pi \mid F(\pi') \geq F(\pi)\}, \quad N := |\Pi|, \quad I(\pi) := -\log_2\left(\frac{|M(\pi)|}{N}\right).$$

Interpretation. A policy has high future function if it is likely to stay correct when new requirements appear. Functional information measures how rare that level of future function is among the policies that are currently viable.

Under exchangeability, $F(\pi)$ is a monotone function of $|B_\pi|$. So inside Π , larger unobserved buffers imply higher functional information. Because α is fixed and $\alpha \subseteq \text{Ext}(\pi)$ for $\pi \in \Pi$, we have $|\text{Ext}(\pi)| = |\alpha| + |B_\pi|$, so this is equivalent to higher weakness up to an additive constant [5].

A Formal Law of Increasing Functional Information

Now consider a population of viable policies sampled uniformly from Π . Selection acts by persistence (i.e. something that does not persist is not selected). A policy survives to the next time step exactly when it survives the novelty set.

Theorem 1 (Law of increasing functional information under novelty selection). *Assume an exchangeable novelty prior P . Draw π uniformly from the viable set Π . Condition on survival, meaning on the event $S \subseteq B_\pi$. Then*

$$\mathbb{E}[I(\pi) \mid \pi \text{ survives}] \geq \mathbb{E}[I(\pi)],$$

with strict inequality whenever $|B_\pi|$ is not constant over Π .

Interpretation. Selection for persistence reweights the population toward policies that keep more untested possibilities open. Those policies are rarer among viable policies. So the average functional information of survivors increases.

A proof is in Supplementary Information [11]. The key step is a one line inequality. When a population is reweighted by a factor that increases with a trait, the conditional mean of any monotone transform of that trait increases.

Experiments in Finite Worlds

We tested the novelty claim in fully enumerated toy worlds where every outcome can be counted. Each world is a finite catalogue L of 48 implementable outcomes. Each outcome is a binary feature vector of length 10.

A current function fixes 2 base commitments that every viable policy must satisfy. We draw 4 observed outcomes α from the outcomes that satisfy those base commitments. Because α is small, some extra feature values look constant by accident. A policy can commit to those accidents and still pass the observed tests. That is overcommitment. It shrinks the unobserved buffer B_π .

We compare two selection rules among policies that satisfy the base commitments and remain compatible with the observed outcomes. The weakness rule chooses a viable policy that maximises $|B_\pi|$. The baseline chooses uniformly at random from the viable set.

Interpretation. Weakness maximisation tries to keep options open. Random does not try.

Under uniform novelty, the counting law gives $\log_2 \mathbb{P}(\text{survive}) = |B_\pi| - |U|$ exactly. So this experiment is noise free. No train test drama, just counting. Figure 1 summarises the mean $\log_2$ survival probability under uniform novelty.

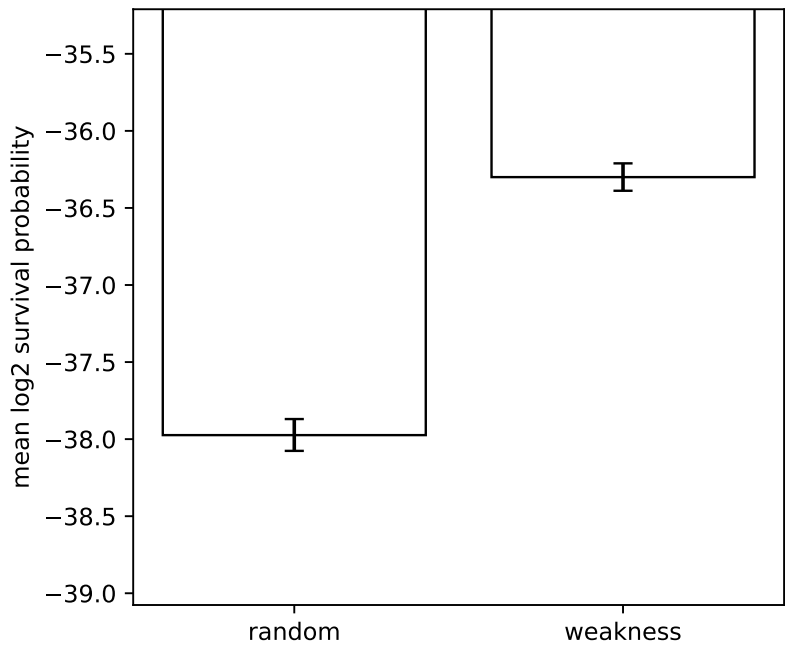


Figure 1. Weakness selection improves survival under novelty. Bars show mean log₂ survival probability under a uniform novelty prior across 4000 trials in 200 toy worlds, with 95% bootstrap confidence intervals. Weakness selection is Bayes optimal in this setting and outperforms random choice by 1.674 bits.

In the uniform novelty setting, weakness selection improves mean log survival probability by 1.674 bits relative to random choice.

Simplicity Avoids Damage and Complexity Needs Repair

The theorem is about novelty selection. Real systems do face both wear and tear. If you have more moving parts, you have more ways to fail. This is where everyday intuition about simplicity is right. A rock survives for boring reasons. It has little to break.

In Stack Theory terms, this is a very different axis from weakness. Weakness counts how many futures you still permit [5]. Simplicity in the physical sense reduces the number of damageable parts. A system can persist by being simple enough that repair is unnecessary for persistence. A system can also persist by being complex and continuously repairing itself. All else being equal a rock that self repairs is more persistent than a rock that does not self repair. However self-repair has a cost in complexity, which makes the self-repairing rock less likely to persist. This divides the world into objects which are simple, and objects which complex but viable because their complexity is balanced by effective self repair. In between these two clusters is *the void of the unviable*. This is a void as described in [12].

Experiment E is a minimal damage model that illustrates this [13]. A system requires k essential components to remain viable. Each component fails independently with probability $p = 0.020$ per time step. Repair can restore at most $r = 1$ failed components per step. We compare a simple system with $k = 3$ and no repair to a complex system with $k = 12$, with and without repair.

Figure 2 shows survival curves.

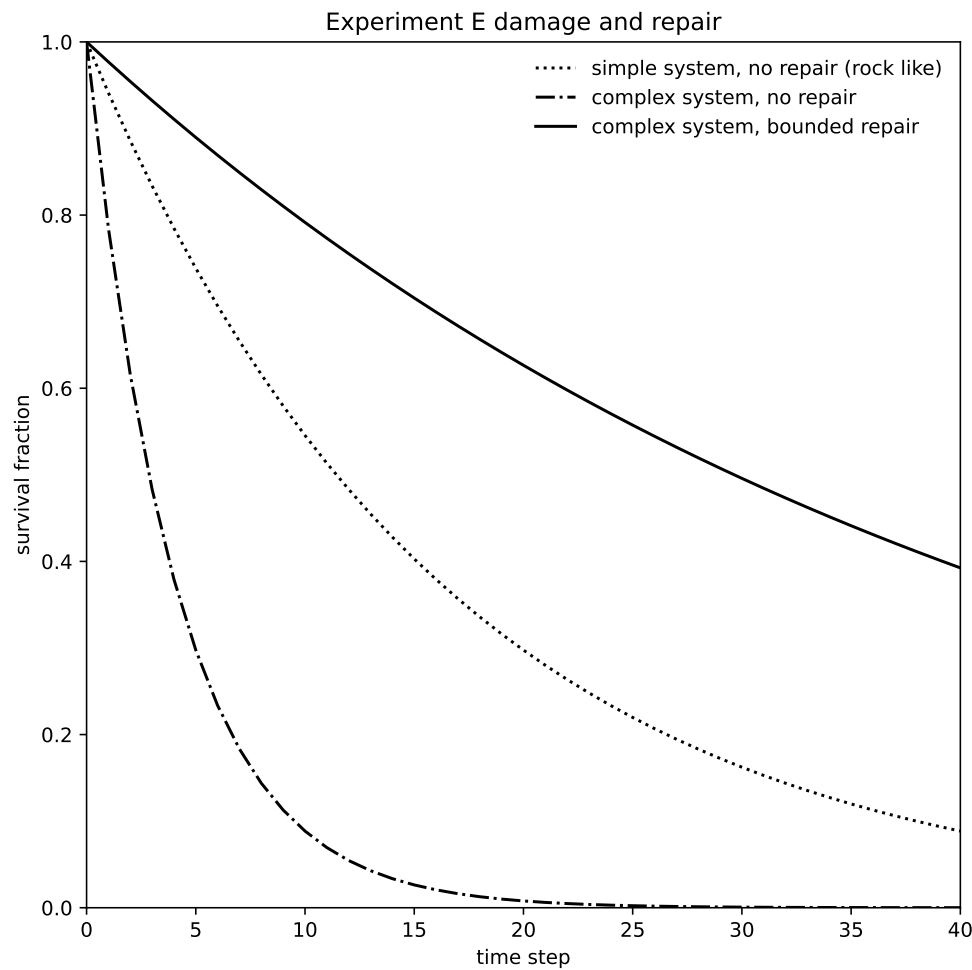


Figure 2. Simplicity avoids damage and complexity requires repair. Survival curves under stochastic damage with bounded repair. The simple system has fewer essential parts so it suffers fewer fatal breaks and needs no repair. The complex system has more to break so it dies quickly without repair. With repair, the complex system becomes viable again.

Interpretation. Simple things persist by not needing repair. Complex things persist by repairing.

Real systems face both novelty and damage. This paper proves a result about novelty selection. The repair model is included only to clarify why simplicity and weakness are different persistence mechanisms.

Discussion

This paper formalises one version of the law of increasing functional information. The law is not a claim that complexity must increase, at least in our interpretation of it [5]. In our interpretation, the law is a claim that selection under novelty biases survivors toward weak constraints that preserve compatibility with many unseen outcomes. This formalises The Cosmic Ought as described in the associated PhD thesis [5].

More broadly, it supports a constraints first view where what persists is shaped by what can be stably realised, not by a monotone complexity gradient [14].

Importantly, simplicity does not explain the novelty result in this framework. Under novelty, the explanatory variable is weakness. Weakness is invariant under re-labellings of the unobserved region, because it is a count of compatible completions [5,15]. Under damage, simplicity can matter because having fewer essential parts reduces the rate at which the world breaks you. That is not a universal simplicity principle. The relationship between robustness and evolvability has been studied extensively in molecular and morphological contexts [16]. Our result adds a distinct channel: weakness

as robustness to novel selection pressures, where compatible completions play the role that neutral neighbours play in sequence space.

The main empirical message is that in finite worlds where we can enumerate all possibilities, weakness maximisation is a reliable proxy for survival under novelty. This aligns with generalisation optimal learning results for weakness in artificial intelligence [17]. When damage and repair become part of the story, we can also see why a single scalar notion of complexity is a trap [13]. Some systems are simple and survive by being hard to break. Other systems are complex and survive by repairing. For example, homeostatic regulation can begin as co-regulation between co-embodied organisms [18], which in our interpretation would align with the idea of a repair budget capable of supporting complexity. Under novelty, the systems that keep their options open are the ones that tend to survive.

Data and Code Availability

Figures 1 and 2 are generated by the Python script `run_all_experiments.py` included with this draft package. The broader code and theory stack is documented in the technical appendices release [11]. A fully reproducible package is included with this manuscript draft.

References

1. Wong, M.L.; Cleland, C.E.; Arend, D.; Bartlett, S.; Cleaves, H.J.; Demarest, H.; Prabhu, A.; Lunine, J.I.; Hazen, R.M. On the roles of function and selection in evolving systems. *Proceedings of the National Academy of Sciences* **2023**, *120*, e2310223120, [<https://www.pnas.org/doi/pdf/10.1073/pnas.2310223120>]. <https://doi.org/10.1073/pnas.2310223120>.
2. Root-Bernstein, M. Evolution is not driven by and toward increasing information and complexity. *Proceedings of the National Academy of Sciences* **2024**, *121*, e2318689121. <https://doi.org/10.1073/pnas.2318689121>.
3. Wong, M.L.; Bartlett, S.; Cleland, C.E.; Demarest, H.; Cleaves, H.J.; Prabhu, A.; Lunine, J.I.; Hazen, R.M. Reply to Root-Bernstein: Increasing complexity allows for the pervasiveness of low-complexity entities and is not anthropocentric. *Proceedings of the National Academy of Sciences* **2024**, *121*, e2406598121. <https://doi.org/10.1073/pnas.2406598121>.
4. Bennett, M.T. Are Biological Systems More Intelligent Than Artificial Intelligence?, 2024, [[arXiv:cs.AI/2405.02325](https://arxiv.org/abs/2405.02325)]. In press, 2026, Philosophical Transactions of the Royal Society B: Biological Sciences. Special issue on Hybrid agencies: crossing borders between biological and artificial worlds, <https://doi.org/10.48550/arxiv.2405.02325>.
5. Bennett, M.T. How To Build Conscious Machines. PhD thesis, The Australian National University, 2025. <https://doi.org/10.25911/ta58-p428>.
6. Kauffman, S.A. *Investigations*; Oxford University Press: Oxford, 2000.
7. Longo, G.; Montévil, M.; Kauffman, S. No entailing laws, but enablement in the evolution of the biosphere. In Proceedings of the Proceedings of the 14th Annual Conference Companion on Genetic and Evolutionary Computation, New York, NY, USA, 2012; GECCO '12, p. 1379–1392. <https://doi.org/10.1145/2330784.2330946>.
8. Wilke, C.O.; Wang, J.L.; Ofria, C.; Lenski, R.E.; Adami, C. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* **2001**, *412*, 331–333.
9. Hazen, R.M.; Griffin, P.L.; Carothers, J.M.; Szostak, J.W. Functional information and the emergence of biocomplexity. *Proceedings of the National Academy of Sciences* **2007**, *104*, 8574–8581. <https://doi.org/10.1073/pnas.0701744104>.
10. Szostak, J.W. Functional information: Molecular messages. *Nature* **2003**, *423*, 689–689. <https://doi.org/10.1038/423689a>.
11. Bennett, M.T. Technical Appendices, 2025. Archived release on Zenodo. Source repository: <https://github.com/ViscousLemming/Technical-Appendices>, <https://doi.org/10.5281/zenodo.17164785>.
12. Solé, R.; Seoane, L.F.; Pla-Mauri, J.; Bennett, M.T.; Hochberg, M.E.; Levin, M. Cognition spaces: natural, artificial, and hybrid, 2026, [[2601.12837](https://arxiv.org/abs/2601.12837)]. <https://doi.org/10.48550/arxiv.2601.12837>.
13. Bennett, M.T. Is Complexity an Illusion? In Proceedings of the 17th International Conference on Artificial General Intelligence. Springer, 2024, Lecture Notes in Computer Science. https://doi.org/10.1007/978-3-031-65572-2_2.

14. Solé, R.; Kempes, C.P.; Corominas-Murtra, B.; De Domenico, M.; Kolchinsky, A.; Lachmann, M.; Libby, E.; Saavedra, S.; Smith, E.; Wolpert, D. Fundamental constraints to the logic of living systems. *Interface Focus* **2024**, *14*, 20240010. <https://doi.org/10.1098/rsfs.2024.0010>.
15. Bennett, M.T. The Optimal Choice of Hypothesis Is the Weakest, Not the Shortest. In Proceedings of the 16th International Conference on Artificial General Intelligence. Springer, 2023, Lecture Notes in Computer Science, pp. 42–51. https://doi.org/10.1007/978-3-031-33469-6_5.
16. Wagner, A. Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B: Biological Sciences* **2007**, *275*, 91–100, [<https://royalsocietypublishing.org/rspb/article-pdf/275/1630/91/593598/rspb.2007.1137.pdf>]. <https://doi.org/10.1098/rspb.2007.1137>.
17. Bennett, M.T. Computational Dualism and Objective Superintelligence. In Proceedings of the 17th International Conference on Artificial General Intelligence. Springer, 2024, Lecture Notes in Computer Science. https://doi.org/10.1007/978-3-031-65572-2_3.
18. Ciaunica, A.; Constant, A.; Preissl, H.; Fotopoulou, K. The first prior: from co-embodiment to co-homeostasis in early life. *Consciousness and cognition* **2021**, *91*, 103117.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.