Author for correspondence:
Amahury Jafet López-Díaz
e-mail: alpez@binghamton.edu

Closing the loop: how semantic closure enables open-ended evolution?

Amahury Jafet López-Díaz and Carlos Gershenson

School of Systems Science and Industrial Engineering, State University of New York at Binghamton, Binghamton, NY, USA

AJL-D, 0009-0004-6315-3474

This study explores the evolutionary emergence of semantic closure—the self-referential mechanism through which symbols actively construct and interpret their own functional contexts—by integrating concepts from relational biology, physical biosemiotics and ecological psychology into a unified computational enactivism framework. By extending Hofmeyr's (Fabrication, Assembly) systems—a continuation of Rosen's (Metabolism, Repair) systems—with a temporal parametrization, we develop a model capable of capturing critical properties of life, including autopoiesis, anticipation and adaptation. Our stepwise model traces the evolution of semantic closure from simple reaction networks that recognize regular languages to self-constructing chemical systems with anticipatory capabilities, identifying self-reference as necessary for robust self-replication and open-ended evolution. Such a computational enactivist perspective underscores the essential necessity of implementing syntax–pragmatic transformations into realizations of life, providing a cohesive theoretical basis for a recently proposed trialectic between autopoiesis, anticipation and adaptation to solve the problem of relevance realization. Thus, our work opens avenues for new models of computation that can better capture the dynamics of life, naturalize agency and cognition and provide fundamental principles underlying biological information processing.

1. Introduction

> 'How, therefore, we must ask, is it possible for us to distinguish the living from the lifeless, if we can describe both conceptually by the motion of inorganic corpuscles?'
>
> Pearson, 1892 [1, p. 341-342]

The question posed by Karl Pearson over a century ago remains central to the philosophical and scientific discourse concerning the nature of life and its differentiation from non-life. Historically, science has attempted to resolve this conundrum through increasingly sophisticated physical and mathematical models. Early attempts, grounded in classical and later quantum physics, explored the possibility of undiscovered laws or unique features intrinsic to living matter [2,3]. The mid-twentieth century explosion in molecular biology and origin-of-life research, culminating in the discovery of DNA's structure and genetic mechanisms, fostered the belief that conventional physical and chemical laws were sufficient to explain life. This view, championed by Kendrew [4], marked the entrenchment of reductionism in theoretical biology. Yet, Pearson's fundamental inquiry persists partially unanswered.

More recently, complex systems science [5–8] has emerged as an alternative paradigm to explain life. Complex systems science seeks universal principles across domains: from spin glasses and sandpiles to cells and societies. According to this narrative, life is an emergent autonomous pattern, produced by the nonlinear interactions among its components. However, complex system science does not discriminate life from non-life in physical or

functional terms. Furthermore, regardless of the many complex systems science models—nonlinear dynamics [9], ergodic theory [10], self-organized criticality [11], agent-based models [12], networks [13] and game theory [14], among others—we have developed, rule-based self-organization can only yield limited novelty, which contrasts drastically with the open-ended evolutionary potential of life to innovate using circumscribed resources.

As a complementary viewpoint to complex systems science, physical biosemiotics (pioneered by Howard Pattee) offers a radically different account of what distinguishes living from non-living systems [15]. Unlike models that seek universal principles across domains through rule-based self-organization or statistical dynamics, Pattee argued that life originates in a fundamental complementarity between symbolic representation and physical dynamics [16]. At the heart of this framework lies the concept of the epistemic cut: a necessary demarcation between quiescent, rate-independent symbolic structures and the rate-dependent, energy-governed dynamics they control [17]. This distinction, which does not correspond to an ontological dualism but an epistemic necessity, reframes the origin of biological order as dependent not only on physical laws but also on the emergence of symbol-mediated constraints capable of controlling those laws. Pattee's work thus sets the stage for a theory of life in which measurement, control and memory are not secondary features but primary requirements for evolvability and functional organization.

This leads us to the concept of symbol. Biological systems are pervaded by codes, from the genetic code to behavioural signals [18,19]. However, symbols do not exist in isolation, but are always part of a semiotic system [3,17]. A semiotic system consists of (i) discrete, rate-independent symbol vehicles; (ii) interpreting structures—non-integrable constraints or codes; and (iii) organizations in which symbols and interpreting structures have functional value [3,20]. These semiotic components conform to the pragmatic theory of communication developed by Morris [21], which distinguishes three irreducible dimensions: syntax (formal relations between symbols), semantics (correspondence with referents) and pragmatics (usefulness or action-invoking roles).

As shown by Cariani [22], the three irreducible semiotic axes are essential in understanding how organisms process and act on information. Still, such an irreducibility challenges our conception of modelling, since we are used to reducing everything to symbols and their syntax. In fact, the whole branch of artificial life is based on this premise [15]. The idea of reducing everything to syntax is not far-fetched. If we identify observables with symbols and their dynamical laws with syntax, then interpreting structures acquire the role of non-holonomic (or non-integrable) constraints: flexible structures that control dynamics without eliminating configurational alternatives. These constraints, such as mechanical governors in machines or enzymes in cells, cannot be separated from the dynamics they influence. Unlike rigid constraints that merely reduce degrees of freedom, non-integrable constraints retain degeneracy, enabling symbolic control without eliminating the system's potential variability. They serve as memory readers that bridge the epistemic cut by coupling quiescent symbols with rate-dependent trajectories [3].

Importantly, constraints are epistemically but not ontologically irreducible. Of course, non-holonomic constraints are theoretically reducible to microscopic dynamics, but this reduction obliterates their functional role. As von Neumann [23] demonstrated, if we describe both a physical system S and its measuring device M within a single dynamical framework, we then require a new device M' to establish the initial conditions for the combined system $S + M$, and so on ad infinitum. This infinite regress reveals a deeper issue: measurement—the act of determining the initial state of a system—cannot be internalized within the same dynamical description without erasing its function.

A measurement device, in this context, is not merely another physical system but a structure that establishes the initial conditions under which dynamical laws operate. Measurement converts a physical state into a symbolic representation, and as such, it plays a role distinct from the dynamics it constrains. Attempting to model both the system and its measurement apparatus within a unified dynamic eliminates the crucial epistemic function of the measuring device as a selector and interpreter of system states. Consequently, measurement—and by extension, control (as the inverse process that translates symbolic instructions into physical outcomes)—cannot be fully captured within a closed dynamical description. The pragmatic function of measurement and control disappears when subsumed into purely energetic or dynamical laws. Thus, the epistemic distinction between symbol and referent, syntax and pragmatics, must be preserved if we are to retain explanatory power [22].

Such a distinction also underpins a critique towards computationalism. Turing's abstract machine [24] manipulates symbols via syntactic rules independent of physical implementation. The illusion that symbol systems are autonomous (detached from time, energy or semantic content) arises from this abstraction. Computation, while powerful for simulating complexity, cannot ground symbolic meaning in physical processes. This isolation is evident in programmable computers, where non-integrable hardware (e.g. switches, memory) allows symbolic control, yet the computer's function appears implementation-free. As Hoffmeyer & Emmeche [25] and Kull [26] note, symbolic expression seems unconstrained by physical law. Yet, this view is misleading: computation is independent of its material implementation only in ideal noiseless, error-free symbol manipulation (27); as Rosen [28] and Emmeche [29] argue, computationalism leads to syntacticalization and derealization, masking the physical substrate of symbolic functions. This insight suggests that capturing the concurrent, adaptive and information-sensitive behaviour observed in living systems [30] requires a departure from classical, Turing-style architectures. Instead, it invites the development of computing paradigms that explicitly incorporate mechanisms of measurement and control [31].

In this way, based on the epistemic cut, Pattee partially answered Pearson's question: *'It is not possible to distinguish the living from the lifeless by the most detailed "motion of inorganic corpuscles" alone. The logic of this answer is that life entails an epistemic cut that is not distinguishable by microscopic (corpuscular) laws. As von Neumann's argument shows, any distinction between subject and object requires a description of the constraints that execute measurement and control processes; and such a functional description is not reducible to the dynamics that is being measured or controlled'* [17, p. 16].

We can think of the genotype–phenotype interface as the primaeval locus of this epistemic cut. Gene sequences are energetically degenerate and not distinguished by physical law; distinctions arise only when proteins fold into active configurations that break degeneracy and initiate function. These distinctions occur only within the context of a self-constructing system, making *semantic*

closure—the self-referential condition where a system's symbols control the construction of its own interpretive mechanisms—necessary for life [27,32]. Semantic closure represents a circularity where genotype informs phenotype, which in turn constructs and interprets the genotype, bridging syntax and pragmatics through self-referential dynamics. However, as Walker & Davies [33] emphasize, the real challenge of life's origin is to understand how those instructional information control systems emerge naturally and spontaneously from mere molecular dynamics. Currently, we lack a general model to explain how such semantic closure originated [27].

This paper explores the origin and early evolution of semantic closure. First, we introduce Rosen's thinking and analyse his (Metabolism, Repair) systems, which are later extended to Hofmeyr's (Fabrication, Assembly) systems. Second, we develop a temporally parametrized extension of (F, A) systems, unifying key properties of life, including autopoiesis, anticipation and adaptation. Third, we elucidate the origin and early evolution of semantic closure by proceeding down the logical chain of assumptions, obtaining weaker versions of our model. Finally, we integrate these findings into a computational enactivist framework consistent with multiple theoretical results that have recently been obtained, showing that there is a self-referential takeover in the evolution of life. Throughout our manuscript, we provide a biologically relevant example of (M, R) system, discuss the realizability of relational models and propose some ways to extend our ideas, including open questions for our research programme and a glossary explaining relevant terms that may be confusing to the reader.

2. A unified representation of life

Throughout the last century, multiple models have been developed to explain the fundamental characteristics of life: the homeostat [34], (M, R) systems [35,36] the kinematic automaton [37], autopoiesis [38,39], dissipative structures [40], autocatalytic sets [41,42], the chemoton [43] and the hypercycle [44], among others. However, although the notion of organizational closure is ubiquitous in all these contemporary frameworks, remarkably, none of them cited others [45], leaving aside a potential unified model where all these conceptualizations turn out to be different faces of the same die. As Korbak [46, p. 40] noted, of all these models, Rosen's (Metabolism, Repair) systems are *'general enough to encompass all kinds of mathematical structures (and formal or computational models)'*, becoming an *epistemological infrastructure* for all modelling in biology [47].

In a long series of papers dating back to 1958 [35,36] and later condensed in the 1990s [48], Robert Rosen introduced the (Metabolism, Repair)-systems paradigm, which was inspired by biotopology [49]. At first, Rosen represented (M, R) systems using graph theory, but he quickly transitioned to category theory, as this shift allowed him to better represent the minimal functional relationships required for life. Historically, Rosen's early papers pre-date the widespread use of the Yoneda lemma, Lawvere's fixed-point theorem and other fundamental category theory results. However, later reconstructions showed (M, R) systems can indeed be derived functorially via Yoneda [50].

Instead, Rosen's adoption of category theory followed Eilenberg–MacLane's original viewpoint: any category embeds faithfully into **Set** (the category of sets), so one may reason with sets without forfeiting structural generality [36,51]. Our aim here is expository and biological: we keep the categorical language to make causes, closure and realizability model-independent, but we avoid imposing extra axioms that would be needed to deploy powerful theorems that are not necessary for the present results.

Let $\mathcal{C}$ be a concrete category, $A, B \in \mathrm{Ob}(\mathcal{C})$, $f \in H(A, B)$ and $\Phi_f \in H(B, H(A, B))$, where $H(X, Y)$ is a proper subset of the hom-set of all mappings from X to Y. According to Rosen [52], by making A, B and f sufficiently complex, it is possible to reduce every (M, R) system to the following diagram:

$$A \xrightarrow{f} B \xrightarrow{\Phi_f} H(A, B). \tag{2.1}$$

Unlike the traditional Newtonian approach, which generally uses state-space-based models with well-defined transition functions, Rosen's approach relies on temporally invariant category theory models, which support multiple modes of causation consistent with Aristotle's four causes. According to Rosen, it is this very segregation into independent categories of causation that prevents the Newtonian picture from describing the entailments and linkage relations characteristic of life [48,53]. In diagram (2.1), A is the material cause of B, f is the efficient cause of B and $H(A, B)$ is the final cause of B. Here, Rosen leaves out the formal cause while making the efficient cause of f (which is Φ_f) internal to the system. At the same time, the efficient cause of Φ_f lies outside the system, requiring an external agent. To resolve this issue, Rosen used the evaluation map of $\mathcal{C}$, which is defined as follows:

$$\mathrm{eval}_{Y,X}^{\mathcal{C}} : Y^X \times X \to Y, \tag{2.2}$$

where Y^X denotes the hom-set of all mappings from X to Y. The evaluation map $\mathrm{eval}_{Y,X}^{\mathcal{C}}$ takes a mapping $f \in Y^X$, an input $x \in X$ and returns the value $f(x) \in Y$. Importantly, given $f \in Y^X$, if we fix a specific $x \in X$ and vary $f \in Y^X$, we can define a new mapping $\hat{x} : Y^X \to Y$, which satisfies $f(x) = \hat{x}(f)$. For Rosen's purposes, the *inverse* evaluation map allows the construction of hierarchical cycles that underlie his (M, R) systems. Building on this foundation, Rosen extended the framework to explicitly incorporate closure. By letting $X = B$ and $Y = H(A, B)$ in equation (2.2), we obtain

$$\hat{b} : H(B, H(A, B)) \to H(A, B).$$

Mapping $\hat{b}$ has a left inverse if it is injective. This means that for every pair of maps $\Phi_1, \Phi_2 \in H(B, H(A, B))$,

$$\hat{b}(\Phi_1) = \hat{b}(\Phi_2) \Rightarrow \Phi_1 = \Phi_2.$$

Equivalently,

$$\Phi_1(b) = \Phi_2(b) \Rightarrow \Phi_1 = \Phi_2. \tag{2.3}$$

Implication (2.3) tells us that if two replacement maps agree at b, then they must agree everywhere. Thus, the replacement map Φ_f is uniquely determined by its one value $\Phi_f(b) = f \in H(A, B)$, endowing closure in (M, R) systems with nothing but the ingredients of metabolism and replacement. Let us denote $\hat{b}^{-1}$ as β_b. Through this construction, Rosen derived a new (M, R) system that incorporates closure, expressed as

$$B \xrightarrow{\Phi_f} H(A, B) \xrightarrow{\beta_b} H(B, H(A, B)).$$ (2.4)

In Rosen's words: '*This construction has two important consequences: (a) it is the only formalism, to my knowledge, which incorporates replication of genetic components as an essential and natural consequence of metabolic activities, and (b) exhibits a new and most interesting formal equivalence between metabolic and replicative aspects in these systems*' [52, p. 52]. By combining the diagrams (2.1) and (2.4), Rosen arrived at the canonical (M, R) system

$$\beta_b \circ \Phi_f \circ f : A \xrightarrow{f} B \xrightarrow{\Phi_f} H(A, B) \xrightarrow{\beta_b} H(B, H(A, B)).$$ (2.5)

In Rosen's [52] terminology, f is the metabolic part (M) of the system (2.5), Φ_f is the repair (R) part of the system (2.5) and β_b is the replicative part of the system (2.5). As Letelier *et al.* [54] explain, it is much more logical from a biological point of view to use the terms 'replacement' and 'closure' rather than 'repair' and 'replication'. Rosen [48] expresses (2.5) as the following diagram:

$$
\begin{array}{ccc}
 & f & \\
A \longrightarrow & B & \longleftarrow \Phi_f
\end{array}
$$ (2.6)

In diagram (2.6), the solid arrows represent the mappings shown in (2.5), and the dashed arrows represent the morphisms f, Φ_f, β_b being applied to their respective input. As the reader could observe, implicitly, Rosen is assuming there is a one-to-one correspondence between B and β_b [55]. This guarantees *closure to efficient causation*. In biological terms, this means that all enzymes, all catalytic processes that are used within the system, need to be synthesized by the organization *per se*, without recourse to any external agent. According to Rosen, this is what characterizes life [48].

2.1. A biological example

In the origins-of-life literature, reflexively autocatalytic and food-generated (RAF) networks [56] and their seed-dependent autocatalytic system (SDAS) variants [45] give concrete instantiations of organizational closure at a chemical level. In RAF/SDAS, a fixed 'food set' together with catalysis constraints selects admissible reaction maps. Seeds trigger higher-level closures that recursively generate their own catalysts. This clearly mirrors the role of $H(A, B)$ (admissible metabolisms) and Φ_f (selector of the executable metabolism) here, while the replication step β_b corresponds to the passage to higher-order closures once a viable catalytic configuration b is present. Thus, our categorical (M, R)-scaffold offers a unifying language that can explicitly host RAF/SDAS chemistries as particular choices of objects, morphisms and admissibility constraints [57].

So far, our treatment of characterizing life has been heavily abstract. To help the reader digest this content and, at the same time, provide a clearer perspective on relational biology, we dedicate this short section to exemplify how we can map Rosen's abstraction to biologically relevant morphisms and objects. First, let us coarse-grain the core gene-expression machinery of a bacterial cell into four catalysts

$$\text{RNAP, Rib, AARS, Chap,}$$

standing, respectively, for RNA polymerase, a ribosome, the (lumped) set of aminoacyl-tRNA synthetases and a chaperone/assembly factor. As *materials* (transformables), we consider pools of precursors

$$\text{AA (amino acids), NTP (nucleoside triphosphates),}$$
$$\text{tRNA, rRNA, mRNA.}$$
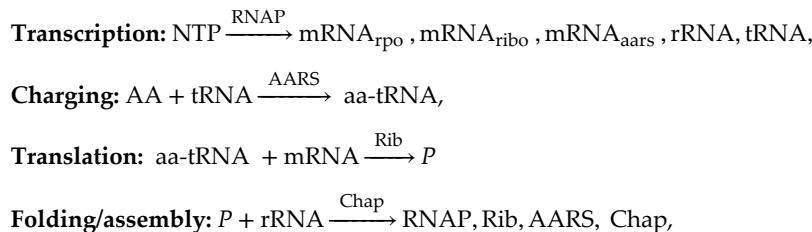
We treat the genome G as a fixed parameter constraining admissible maps. The objects are now specified as *sets of admissible configurations* and their Cartesian products

$$A := \text{AA} \times \text{NTP} \times \text{tRNA} \times \text{rRNA} \times \text{mRNA},$$

and

$$B := \text{RNAP} \times \text{Rib} \times \text{AARS} \times \text{Chap}.$$

Therefore, an element $a \in A$ gives material levels/compositions, while $b \in B$ is a 4-tuple of catalytic states. With this, metabolism $f \in H(A, B)$ is the composite action of the following schematic reactions (arrows labelled by their catalysts):

Transcription: $\text{NTP} \xrightarrow{\text{RNAP}} \text{mRNA}_{\text{rpo}}, \text{mRNA}_{\text{ribo}}, \text{mRNA}_{\text{aars}}, \text{rRNA}, \text{tRNA},$

Charging: $\text{AA} + \text{tRNA} \xrightarrow{\text{AARS}} \text{aa-tRNA},$

Translation: $\text{aa-tRNA} + \text{mRNA} \xrightarrow{\text{Rib}} P$

Folding/assembly: $P + \text{rRNA} \xrightarrow{\text{Chap}} \text{RNAP}, \text{Rib}, \text{AARS}, \text{Chap},$

where P represents the pool of unfolded polypeptides, including RNAP/Rib/AARS/Chap chains. At the level of $\mathcal{C}$, this induces a time-invariant morphism

$$f : A \longrightarrow B, \quad a \longmapsto b = f(a).$$

Let $H(A, B)$ be the set of *admissible* metabolic maps compatible with (i) the fixed genome G, (ii) mass-balance/assembly constraints, and (iii) coarse-grained kinetics. Given a catalytic configuration $b \in B$ (current RNAP/Rib/AARS/Chap states), the *selector/replacement* map

$$\Phi_f \in H\big(B, H(A, B)\big)$$

returns the unique admissible metabolism $f_b \in H(A, B)$ that *this very configuration* can realize on genome G

$$\Phi_f(b) = f_b, \quad \text{with } f_b(a) = b' \in B.$$

Intuitively, the particular RNAP holoenzyme, ribosome stoichiometry and AARS repertoire determine which $A \to B$ map is executable (promoter usage, translation capacity, assembly constraints). Now, let us pick a *distinguished* self-consistent configuration

$$b^{\star} = \big(\text{RNAP}_{\sigma}, \text{Rib}_{70S}, \text{AARS}_{\text{complete}}, \text{Chap}_{*}\big) \in B,$$

where $\text{RNAP}_{\sigma} \in \text{RNAP}$ is a housekeeping σ-holoenzyme, $\text{Rib}_{70S} \in \text{Rib}$ is an assembled ribosome, $\text{AARS}_{\text{complete}} \in \text{AARS}$ is a complete minimal set and $\text{Chap}_{*} \in \text{Chap}$ is an adequate chaperone pool. We can restrict the selector space to those admissible under G

$$S(G) := H\big(B, H(A, B)\big).$$

$S(G)$ contains selectors that (i) read the same genome G, (ii) preserve the assembly/stoichiometry rules, and (iii) differ, if at all, only by regulation consistent with $b^{\star}$. On this reduced $S(G)$, we assume an injective evaluation at $b^{\star}$

$$\forall \Phi_1, \Phi_2 \in S(G) : \Phi_1(b^{\star}) = \Phi_2(b^{\star}) \implies \Phi_1 = \Phi_2.$$

Equivalently, the evaluation map

$$\text{eval}_{A,B}^{\mathcal{C}} : S(G) \longrightarrow H(A, B), \quad \Phi \longmapsto \Phi(b^{\star}),$$

is one-to-one. Hence, it admits a left inverse on its image

$$\beta_{b^{\star}} := {\text{eval}_{Y,X}^{\mathcal{C}}}^{-1} : H(A, B) \longrightarrow S(G) \subseteq H\big(B, H(A, B)\big),$$

so that $\beta_{b^{\star}}(f) = \Phi_f$ and $\Phi_f(b^{\star}) = f$. In this way, the (M, R)-loop is realized by the gene-expression core under G, with *objects* of $\mathcal{C}$ read as material/catalyst pools and *morphisms* as genome- and assembly-constrained transformations between them. This same reading also lets RAF/SDAS early metabolisms serve as worked examples: A is a food-plus-intermediate pool, B is the active catalyst set, $H(A, B)$ is the reaction-closure maps permitted by catalysis and availability, Φ_f selects which closure a given $b \in B$ can realize, and β_b records how one viable configuration pins down the entire selector.

2.2. Realizing relational models

Rosen's assertion that no model of closure to efficient causation could be Turing-simulable has often been misunderstood or misrepresented in the literature [58]. As emphasized by Cornish-Bowden & Cárdenas [45], many critiques of Rosen's claim stem from a failure to grasp the core logic of relational biology. This is entirely understandable, as it is a fact that Rosen's ideas changed abruptly over time [59]. For example, Rosen himself investigated the representation of relations, such as (2.1), using sequential machines, extensively exploring the implications of this approach [60–62]. It was the inclusion of replication mechanisms, such as those represented by (2.4), that posed challenges [50].

Later, in what is possibly his best-known work [48], he never used the term *computable*. Instead, Rosen referred to machines, mechanisms and the notion of simulation. Nevertheless, we can get an idea of his late notion of computability from there:

'Thus, the word "simulable" becomes synonymous with 'evaluable by a Turing machine'. In the picturesque language of Turing machines, this means the following: if f is simulable, then there is a Turing machine T such that, for any word w in the domain of f, suitably inscribed on

an input tape to T, and for a suitably chosen initial state of KD-TILDE-PLACEHOLDER-KD T, the machine will halt after a finite number of steps, with f(w) on its output tape' [48, p. 192].

From the above, a myriad of interpretations has arisen, even calling into question the validity of Rosen's arguments [55]. Despite this, it is a fact that Rosen's reasoning lies in the impossibility of capturing self-reference by means of sequential models of computation. In the same way that a Turing machine cannot decide whether it halts on its own description (the Halting Problem), a single automaton cannot decide its own entailment structure when the entailments depend on itself in a circular manner. For example, if f is produced by B, but f is also the catalyst that turns A into B, then we end up needing f to define f. This is a classic impredicative loop [63]. Thus, even if we allow dynamic label reassignment, that does not help with the causal loop unless the automaton can change its own state transition structure during computation based on its own internal outputs. That would require meta-computation, something not allowed in classical computer science.

Using λ-calculus to represent relational models is theoretically appealing [64]. Nevertheless, as shown by Cárdenas *et al.* [55], the mathematical solutions it provides may not directly correspond to biological phenomena, necessitating further refinement and interpretation. A promising resolution to this challenge is provided by Palmer *et al.* [65], who propose that while a single automaton cannot model an organism, a set of automata communicating and sharing information might achieve this. It is clear that Palmer *et al.* [65] wanted to create a *model* of Rosen's replicative (M, R) system rather than just another simulation [64,66,67]. However, at the end of their manuscript, they clarified that their communicating X-machine model does not, as the modelling relation requires, map to the (M, R)-system relations [65]. This creates even more obstacles if we want to *model* (not simulate) fabrication, self-assembly and an actively maintained assembly-enabling environment. As we shall see later, these three aspects are essential to expand Rosen's ideas.

As Pattee [15] points out, simulation accuracy does not translate into effective functional realizations. Most importantly, simulations should never be conflated with realizations. When we simulate life, we get rid of its semantic and pragmatic aspects, focusing merely on its syntactic nature. Simulations—and more generally, consistent formal systems (i.e. not allowing contradictions)—require a predefined set of rules in which novelty is limited. Realizations of life, on the other hand, are capable of generating novelty by interacting directly with their environment, capturing with it those semantic-pragmatic aspects that simulation lacks. This distinction is critical because it underscores the limitations of purely algorithmic models in fully capturing the self-referential and anticipatory nature inherent in biological systems [68]. We have characterized computation as transformations from symbols to symbols, emphasizing that both measurement and control are fundamentally irreducible to computation [23]. Thus, if instead of merely simulating, we construct realizations of relational models, we can inherently incorporate measurement and control into their architecture, becoming capable of interacting with and adapting to their environment. This blueprint effectively transcends the purely symbolic definition of computation we initially presented and allows us to better capture the essence of life [22].

Another relevant aspect when building life realizations is its inherent stochasticity [30]. At the molecular origin of life, stochasticity manifested significantly through ambiguity in early genetic codes and high translational error rates, resulting in heterogeneous peptide populations, called *statistical proteins* [19]. These statistical proteins acted as transitional molecular forms that allowed primitive biological systems to broadly explore chemical and functional spaces, potentially providing adaptive advantages and evolutionary pathways toward increased molecular specificity and coding accuracy. Therefore, regardless of its implementation substrate, any realization of life must be intrinsically stochastic.

2.3. (Fabrication, Assembly) systems

(M, R) systems face two significant challenges. First, they depend on injectivity in $H(A, B)$ to ensure closure, a condition that imposes strict mathematical constraints for the concrete category $\mathcal{C}$ [57,69]. Second, the one-gene-one-enzyme hypothesis that originally inspired Rosen's model is now recognized as overly simplistic [70]. Thus, Rosen's model is not biologically sound. A promising alternative that addresses these limitations is provided by (Fabrication, Assembly) systems [71]. These systems extend Rosen's approach by incorporating free-standing formal causes, allowing for the modelling of both physiological and evolutionary adaptation.

To proceed, it is necessary to clarify what we mean by formal cause. As Hofmeyr [72] explains, formal causation can be integrated with Aristotle's other causes (material and efficient) in multiple ways. Broadly, formal causes can be either free-standing or intrinsic. In the free-standing case, the formal cause exists as an independently specified sequence of rules that directs either the material or efficient cause in carrying out a transformation towards final cause. Examples include a catalyst-template system where monomers are assembled into polymers, or tagging enzymes that act on target-tag pairs to produce a tagged target. In the first example, the enzyme (efficient cause) works in tandem with a separate template (formal cause); in the second, the material cause (target) interacts with an independent tag acting as the formal cause.

By contrast, intrinsic formal causation occurs when the formal cause is embedded within the structure or properties of the material or efficient cause itself. Examples include enzyme active sites with specific substrate-binding properties and unfolded polymers folding into functional conformations under suitable conditions. In the first case, the enzyme active site is an informed efficient cause—it can act only on specific substrates. In the second, the monomer sequence of the polymer acts as an informed material cause—it folds predictably when placed in a particular physical context. After this clarification, we can now show the simplest representation of an (F, A) system,

$$A \xrightarrow{f} B \xrightarrow{\Phi_f} H(A, B), \tag{2.7}$$

$$D \xrightarrow{f} H(B, H(A, B)), \tag{2.8}$$

where $f = I \times \{f_1\} + f_2$, $I = \{I_1, I_2\}$ and $B = B_1 + B_2$. Here, the addition symbol is not a simple algebraic sum but rather indicates that f_2 is being incorporated into the overall process. In (F, A) systems, f has a dual role: f_1, guided by free-standing formal causes (I), produces B, while f_2, an informed efficient cause, produces Φ_f. This allows the system to remain closed to efficient causation while being open to formal causation [72].

From a biosemiotic perspective, this structure becomes especially revealing. As Cariani [22] points out, three of Aristotle's causes align naturally with the semiotic triad identified by Morris [21]: syntax corresponds to the free-standing formal cause, semantics to the efficient cause, and pragmatics to the final cause. Furthermore, material cause translates as symbol vehicle (substrate), informed efficient causes are semantic interpreters whose specificity is embedded within them, and informed material causes are symbolic substrates whose syntax is materially inscribed. This translation is highly suggestive. If (F, A) systems are closed to efficient causation but open to free-standing formal causation, they fundamentally exhibit semantic closure and possess adaptive syntax. The latter implies that the system is no longer constrained by a fixed rule set, thus enabling the representation of affordances by Kauffman & Roli [73].

As a result, (F, A) systems achieve organizational closure and heritable variability while circumventing the need for injectivity in $H(A, B)$ as a prerequisite for closure, thereby rendering the model more biologically realistic. Perhaps a much more intuitive way to understand the above is to look at the diagram that represents this class of models,

$$\tag{2.9}$$

As shown by Hofmeyr [71], (F, A) systems allow the interpretation of concepts such as the metabolism and repair components of Rosen's (M, R) system [36], John von Neumann's kinematic automaton [37], Howard Pattee's symbol-function split via the symbol-folding transformation [74], Marcello Barbieri's genotype–ribotype–phenotype ontology [75] and Tibor Gánti's chemoton [43]. Thus, (F, A) systems are endowed with features that characterized life, such as thermodynamic openness, catalytic specificity, catalytic closure, structural closure, information handling, reproduction and evolvability [45].

In this way, (F, A) systems provide a canonical framework for representing biological organization, demonstrating how hierarchical cycles cannot be fully captured by algorithmic models due to their intrinsic collective impredicativity [68]. For further details on how (F, A) systems can represent the logic underlying a self-manufacturing cell, we recommend consulting Hofmeyr [71]. In the next section, we demonstrate how (F, A) systems can be temporally extended, naturally giving rise to anticipatory systems. This integration allows us to build a computational enactivist narrative that resolves the limitations of algorithmic approaches to modelling anticipation in non-human and non-cognitive living systems.
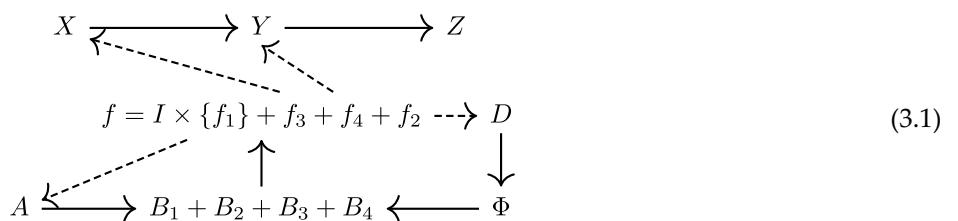
3. Extending (Fabrication, Assembly) systems

One limitation of the current (F, A) system model is its blindness to the external environment. While it accounts for internal self-maintenance, it does not explicitly incorporate mechanisms for environmental sensing and response, both of which are essential for biological measurement (represented by perception carried out by sense organs) and control (represented by actions carried out by various effectors, such as motor and secretory organs) [22]. Notwithstanding, Hofmeyr notes that this limitation can be addressed by incorporating membrane receptors and signal transduction networks into the fabrication machinery [71]. Because these components are encoded in DNA, they can be naturally integrated into the (F, A)-system manufacturing architecture. This insight highlights the importance of extending the model to capture adaptive responses, further enhancing its biological realism from an ecological point of view.

Let $I = \{I_1, I_2, I_3, I_4\}$ be an extended set of free-standing formal causes, where I_3 and I_4 specifically represent the instructions for manufacturing the sensing and response architectures, respectively. Associated, respectively, with I_3 and I_4, we have $f_3 : X \to Y$ (membrane receptors) and $f_4 : Y \to Z$ (signal transducers) for external signals X, internal signals Y and intracellular responses Z. These mappings are such that

$$X \xrightarrow{f_3} Y \xrightarrow{f_4} Z.$$

Notice that these mappings do not form part of covalent metabolic chemistry. Thus, we can extend diagram (2.9) as follows:

$$\tag{3.1}$$

In diagram (3.1), B_3 and B_4 are the components from which membrane transporters and signal transducers are assembled (their unfolded polypeptides). The use of free-standing formal causes (I_3, I_4) to encode membrane receptors and signalling networks opens a door to embedding environmental affordances [76]. In this context, affordances are not fixed properties of the world but *transjective*, relational possibilities for action that emerge through an organism's embodied engagement with its environment. This perspective, emphasized in both ecological psychology [76] and evolutionary accounts of agency [68], highlights the role of the agent in delimiting a meaningful 'arena' of possible interactions. Similar to the view defended in Kauffman & Roli [73], affordances are neither reducible to observer-independent features nor internal subjective projections; rather, they are *enacted* through the dynamic, co-constitutive relation between organism and environment. Incorporating this framework within the extended (F, A) system highlights its capacity not just for self-maintenance but also for situated adaptive behaviour.

Biological systems, however, adapt across multiple time scales, including the rapid biochemical reactions of molecular processes, short-term learning, development and evolution. These temporal scales are critical to understanding how short-term dynamics influence long-term stability and vice versa [77]. To capture this interplay, we associate distinct time scales with each process within the (F, A) system. Importantly, this approach aligns with the notion of constraint developed by Montévil & Mossio [77], which is equivalent to Rosen's concept of efficient causation. We start by rewriting the fundamental relations in our extended (F, A) systems,

$$A \xrightarrow{f} B \xrightarrow{\Phi_f} H(A, B),$$
$$D \xrightarrow{f} H(B, H(A, B)),$$
$$X \xrightarrow{f_3} Y \xrightarrow{f_4} Z,$$

(3.2)

where $f = I \times \{f_1\} + f_2 + f_3 + f_4$, $I = \{I_1, I_2, I_3, I_4\}$ and $B = B_1 + B_2 + B_3 + B_4$. Each arrow shown in (3.2) represents a process, and thus it will have a characteristic temporal scale. For simplicity, let $C = H(A, B)$ and $E = H(B, H(A, B))$. To emphasize the multifunctional role of f, we can rename the different mappings of (3.2) in the following way:

$$A \xrightarrow{\phi} B \xrightarrow{\psi} C$$
$$D \xrightarrow{\theta} E$$
$$X \xrightarrow{\alpha} Y \xrightarrow{\beta} Z.$$

(3.3)

Thus, we assign a time scale τ_x to each process $x \in \{\phi, \psi, \theta, \alpha, \beta\}$. The mappings in (3.3) can now be rewritten as

$$A \xrightarrow[\tau_\phi]{\phi} B \xrightarrow[\tau_\psi]{\psi} C$$
$$D \xrightarrow[\tau_\theta]{\theta} E$$
$$X \xrightarrow[\tau_\alpha]{\alpha} Y \xrightarrow[\tau_\beta]{\beta} Z.$$

(3.4)

Following Montévil & Mossio [77], two scales must be considered for every efficient causation included in a closed system: one scale at which the efficient causation is associated with a time symmetry and another at which it is produced and/or maintained. Let $\Delta_\phi = \tau_\phi - \tau_\psi$, $\Delta_\psi = \tau_\psi - \tau_\theta$, $\Delta_\theta = \tau_\theta - \tau_\psi$, $\Delta_\alpha = \tau_\alpha - \tau_\psi$ and $\Delta_\beta = \tau_\beta - \tau_\psi$. Since the system is closed to efficient causation, at least one Δ_i must be positive, and another Δ_j ($j \neq i$) must be negative [77]. This maintains the temporal coherence of the organization, allowing for organizational closure as the system unfolds over time. Using the methodology proposed by Korbak [46], we represent the temporal parametrization of (3.4) as the asynchronous dynamic Bayesian network shown in figure 1.

In figure 1, $\tau = \tau(\tau_\phi, \tau_\psi, \tau_\theta, \tau_\alpha, \tau_\beta)$, allowing us to model multiple parallel dependencies and interactions probabilistically. Following Korbak [46], we reinterpret each component and each solid line as a random variable, while reinterpreting each dashed line as a dependency. The corresponding probabilistic graphical model is shown in figure 2. Thus, the simulation of an (F, A) system reduces to inferring the following joint distribution:

$$p(\phi_\tau, \psi_\tau, \theta_\tau, \alpha_\tau, \beta_\tau, A_\tau, B_\tau, C_\tau, D_\tau, E_\tau, X_\tau, Y_\tau, Z_\tau, \dots$$
$$\phi_T, \psi_T, \theta_T, \alpha_T, \beta_T, A_T, B_T, C_T, D_T, E_T, X_T, Y_T, Z_T) =$$
$$\prod_{t \geq 0}^{T-\tau} [p(A_{t+\tau}) p(\phi_{t+\tau} \mid A_{t+\tau}, C_t) p(B_{t+\tau} \mid \phi_{t+\tau}) p(\psi_{t+\tau} \mid B_{t+\tau}, E_t)$$
$$p(C_{t+\tau} \mid \psi_{t+\tau}) p(D_{t+\tau}) p(\theta_{t+\tau} \mid D_{t+\tau}, C_t) p(E_{t+\tau} \mid \theta_{t+\tau}) p(X_{t+\tau})$$
$$p(\alpha_{t+\tau} \mid X_{t+\tau}, C_t) p(Y_{t+\tau} \mid \alpha_{t+\tau}) p(\beta_{t+\tau} \mid Y_{t+\tau}, C_t) p(Z_{t+\tau} \mid \beta_{t+\tau})].$$

(3.5)

Through ancestral sampling, we can infer this distribution [46]. Moreover, this temporal parametrization links (F, A) systems to active inference, conceptualizing primitive living systems as generative models probabilistically tailoring their internal states and environments. Since the concept of Markov blanket [78] is compatible with embodied, enactive and extended conceptions of
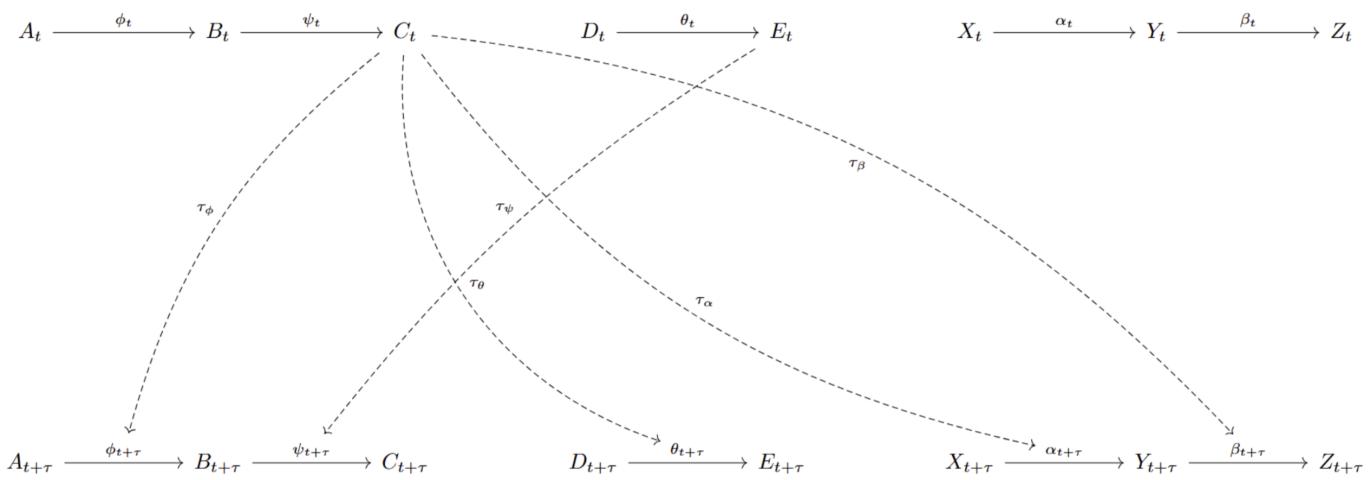
Figure 1. Temporal parametrization of extended (F, A) systems.
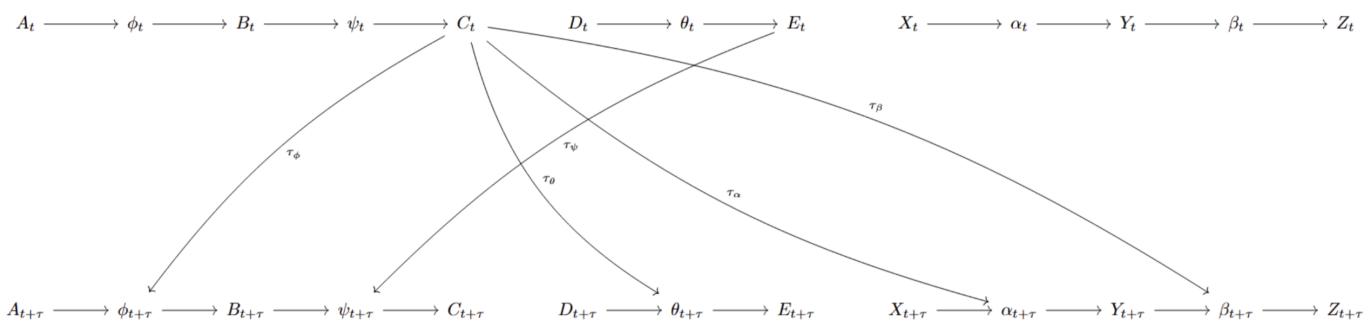
Figure 2. Probabilistic graphical model associated with the temporal parametrization of extended (F, A) systems.
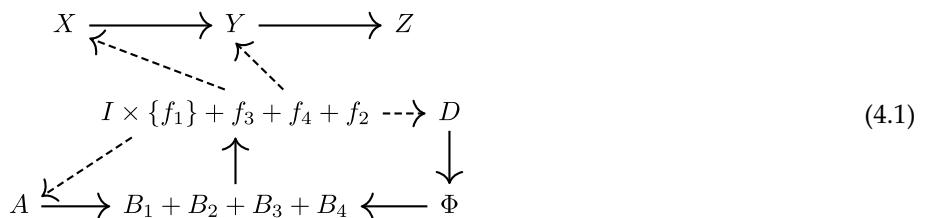
cognition [79], it is possible to construct such a dependency structure for (F, A) systems. Such a construction is potentially compatible with other recent models that have been proposed to understand agency [80–82]. These detailed explorations are beyond the scope of this paper.

Here, it suffices to note that (F, A) systems inherently embody a form of anticipation. By integrating mechanisms for sensing and response through free-standing formal causes and embedding them within asynchronous, multi-scale temporal structures, the system develops the capacity to interpret environmental inputs before they fully unfold. This anticipatory structure resonates with Rosen's conception of anticipatory systems as those that contain internal models capable of projecting future states and acting upon those projections in the present [53].

In Rosen's formal terms, such systems exhibit a predictive model whose outputs serve not merely as representations, but as causal drivers of current behaviour, apparently violating classical notions of causality, yet foundational for biological adaptivity [53]. Our temporally extended (F, A) systems instantiate this idea constructively: rather than relying on retrospective feedback, the system deploys present environmental cues to actively constrain future states through model-mediated inference. This marks a shift from mere reaction to primitive interpretation. Since we are showing that a genuine process of interpretation takes place within simple life forms, such as a cell, our extension of (F, A) systems reconciles Barbieri's code biology [83] with Pattee's physical biosemiotics [17]. The following section is devoted to the study of weak versions of the relational model developed thus far to explore the origin and evolution of semantic closure.
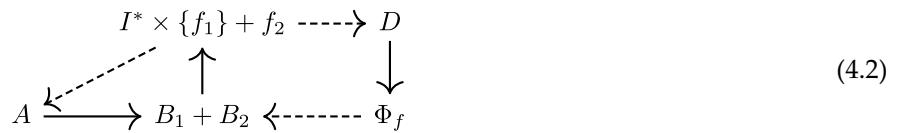
4. Deriving weaker models for primitive semantic closure

From our previous analysis, we know that diagram (4.1) accounts for semantic closure. Having found a minimal model with this feature, we will now derive weaker versions of diagram (4.1) to glimpse the evolutionary emergence of semantic closure.

$$(4.1)$$
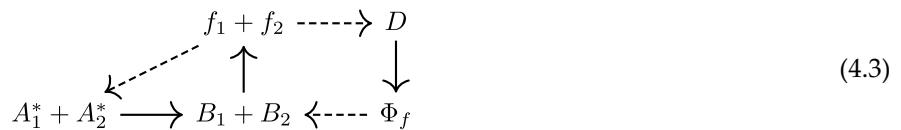
Pattee [84] discusses how hierarchical levels of control in biology emerge through evolution. In early life, simple functional switches (e.g. enzyme activation or inhibition) would allow a system to respond to internal changes, manage its own metabolic

processes and gradually evolve towards more complex forms of environmental interaction, where sensing and adaptation to external stimuli would be added later. Thus, our first step to proceed down our logical chain of assumptions and produce a weak version of the diagram (4.1) is to eliminate $f_3 + f_4$ and $B_3 + B_4$, obtaining the diagram (4.2) shown below.

$$
\begin{array}{ccc}
I^* \times \{f_1\} + f_2 & \dashrightarrow & D \\
\nwarrow \quad \uparrow & & \downarrow \\
A \xrightarrow{\quad} B_1 + B_2 & \dashleftarrow & \Phi_f
\end{array}
\tag{4.2}
$$
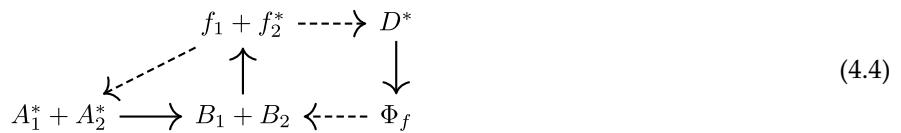
In diagram (4.2), $I^* = \{I_1, I_2\}$. Although digital information processing is a fundamental requirement for life [37,85,86], according to Pattee [84], digital messengers emerge as a product of a synergistic process, akin to how formal rules of syntax and dictionaries evolved from the functional use of primitive symbols in a complex environment. Along these lines, digitization would have arisen as a natural outcome of the processes in reaction networks that had once been primarily analogue. This implies that $I^* = \{I_1, I_2\}$ in diagram (4.2) could not have existed at the beginning.

Since early chemical systems relied only on geophysical and geochemical constraints, with no external formal causes or instructions [84], there are two possibilities to resolve infinite regress [72,87]. The first is that f_1 is an *informed* efficient cause (a semantic interpreter whose specificity is embedded in it), as f_2 already is. The second is that A, although primarily a material cause, assumes the role of an *informed* material cause (a symbolic substrate whose syntax is materially inscribed), similar to B_1 and B_2 [71]. The second option represents the most plausible evolutionary pathway for implementing a weaker version of our (F, A) system [87]. This leads us to the successor diagram below.

$$
\begin{array}{ccc}
f_1 + f_2 & \dashrightarrow & D \\
\nwarrow \quad \uparrow & & \downarrow \\
A_1^* + A_2^* \xrightarrow{\quad} B_1 + B_2 & \dashleftarrow & \Phi_f
\end{array}
\tag{4.3}
$$

In diagram (4.3), A_1^* and A_2^* represent informed material causes, f_1 maps the disjoint set $A_1^* + A_2^*$ to the disjoint set $B1 + B2$, and f_2 maps D to Φ_f. This evolutionary pathway relies on intrinsic constraints and properties rather than complex external mechanisms. Prior to this stage, however, f_2 must be replaced by a non-informed efficient cause (f_2^*), and D must transition into an informed material cause (D^*).

$$
\begin{array}{ccc}
f_1 + f_2^* & \dashrightarrow & D^* \\
\nwarrow \quad \uparrow & & \downarrow \\
A_1^* + A_2^* \xrightarrow{\quad} B_1 + B_2 & \dashleftarrow & \Phi_f
\end{array}
\tag{4.4}
$$

In diagram (4.4), D^* acts as an informed material cause that is prone to being transformed by f_2^* into Φ_f. This architecture retains the autonomy and hierarchical closure of the system. If $\Phi_f = \Phi_{f_1} + \Phi_{f_2}$ and B_2 is provided by the environment (and therefore does not have to be made from A_2^*), then diagram (4.4) can be split into two components [71]. The first component is

$$
\begin{array}{ccc}
 & f_1 & \\
\nwarrow & \uparrow & \\
A_1^* \xrightarrow{\quad} B_1 & \dashleftarrow & \Phi_{f_1}
\end{array}
\tag{4.5}
$$

Although this component is not closed to efficient causation, given that it has the form of a non-replicating (M, R) system, the states and inputs cannot be separated, so it cannot be instantiated as a finite-state automaton [60]. We can represent diagram (4.5) as the following hierarchical cycle:

$$
\begin{array}{ccc}
 & & \Phi_{f_1} \\
 & \swarrow & \\
B_1 & \xrightarrow{\quad} & f_1 \\
 & \swarrow & \\
A_1^* & \xrightarrow{\quad} & B_1
\end{array}
\tag{4.6}
$$

As Palmer *et al.* [65] explained, such a hierarchical cycle can be represented by an automaton with memory. We can then further split the diagram (diagram (4.6)) into

$$\begin{array}{c} \Phi_{f_1} \\ \nwarrow \\ B_1 \longrightarrow f_1 \end{array} \qquad (4.7)$$

and

$$\begin{array}{c} f_1 \\ \nwarrow \\ A_1^* \longrightarrow B_1 \end{array} \qquad (4.8)$$

Both diagrams (4.7) and (4.8) can be represented using a finite-state automaton [60,65]. The second component of diagram (4.4) is

$$\begin{array}{ccc} f_2^* & \dashrightarrow & D^* \\ \uparrow & & \downarrow \\ B_2 & \dashleftarrow & \Phi_{f_2} \end{array} \qquad (4.9)$$

Unlike the first, this component is still closed to efficient causation and thus semantically closed. A rather pertinent observation is that there is a clear parallel between diagram (4.9) and the *autogenic virus* ('autogen' for short) proposed by Deacon [88]. Autogen arises from the interaction of two self-organizing processes: *reciprocal catalysis* and *self-assembly*, the former producing locally asymmetrically high concentrations of a small number of molecular species and the latter requiring persistently high local concentrations of a single species of component molecules. Autogen's molecular logic is summarized by diagram (4.10).

$$\begin{array}{ccc} f + g & \dashrightarrow & a \\ \uparrow & & \downarrow \\ d & \dashleftarrow & n + c \end{array} \qquad (4.10)$$

In diagram (4.10), a and d are material substrates provided by the environment. In addition, f and c are molecular catalysts, while n and g are by-products of the chemical interaction, where g promotes self-assembly [88]. Since in diagram (4.9) f_2^* represents the maintenance of the intracellular milieu Φ_{f_2} that makes self-assembly possible [71], we can equate $f + g \triangleq \Phi_{f_2}$. Thus, $n + c \triangleq f_2^*$, $a \triangleq B_2$ and $d \triangleq D^*$. In this way, Deacon's [88] autogen is a particular case of the model presented here. Returning to diagram (4.9), before achieving closure to efficient causation, the chemical reaction network must have been open. This can be represented as

$$\begin{array}{c} \Phi_{f_2} \\ \nwarrow \\ B_2 \longrightarrow f_2^* \\ \nwarrow \\ D^* \longrightarrow E \end{array} \qquad (4.11)$$

or

$$\begin{array}{c} f_2^* \\ \nwarrow \\ D^* \longrightarrow \Phi_{f_2} \\ \nwarrow \\ B_2 \longrightarrow E \end{array} \qquad (4.12)$$

In diagrams (4.11) and (4.12), E simply represents another chemical product. In both cases, this system is no longer closed to efficient causation, allowing its hierarchical entailments to be modelled using a stream X-machine [65]. When the functional entailments are further dismantled, simpler components acquired the form (4.7)–(4.8) and can be represented by finite-state automata [60]. From the above, we can glimpse a clear transition across the computational capabilities from simple chemical reactions to quasi-autonomous organisms, such as our initial extended (F, A) system. Based on our theoretical observations, in the next section,

we propose an evolutionary narrative in computational enactivist terms [89] for the origin of semantic closure. As we shall see, such an account is consistent with the experimental results obtained by Dueñas-Díez & Pérez-Mercader [90].

5. Towards a computational enactivist narrative for the origin and early evolution of semantic closure

Following the line of thought proposed by Pattee [84], the primordial ecosystem can be conceptualized as the first information-containing architecture on Earth, where primitive geochemical cycles gave rise to a well-defined syntax (based on chemistry) spontaneously and without requiring genetic instructions or initial metabolic control. Given a set of primordial geophysical and geochemical constraints—a primaeval *ecosystem language* [84]—we can envisage the emergence of simple bimolecular reactions, such as those represented in diagrams (4.7) and (4.8), which can be realized as finite-state automata. These simple chemical reactions are capable of recognizing regular languages, a task that does not require counting or memory [90].

In this primitive world, molecules proficient in catalysing key environmental reactions and processes became abundant over time [91]. Later, the diversity of such molecules allowed interactions and, in some cases, bonding, leading to systems of multiple interconnected reactions mediated by common intermediate chemical species. In this panorama, some intermediates functioned both as the product of one reaction and as the reactant for another, forming networks of increasingly complex chemical reactions. These ideas align with the emergence of composomes [92,93] and their maintenance [94].

Such complex reactions can be represented through hierarchical functional entailments that are not closed to efficient causation (see diagrams (4.6), (4.11) and (4.12)) and can be realized using stream X-machines, as suggested by Palmer *et al.* [65]. This primitive form of memory, based on the reuse of catalytic elements, enables the recognition of more complex sequences of information [95] and enhances the evolutionary robustness of such information under changing environmental conditions [33]. Despite their simplicity, these chemical reactions could encode correlations in environmental variables [96] and potentially exhibit a basic form of anticipation [97].

This primaeval landscape helps outline the evolutionary trialectic proposed by Jaeger *et al.* [68], where the environment plays a crucial role in shaping the catalytic abilities of chemical species via random mutations induced by geochemical and geophysical processes. This activity fosters molecular Darwinism [91], where the regulation of terrestrial cycles and adaptation to abrupt changes become, respectively, the goals and actions of this transjective trialectic. Furthermore, this competition among chemical components also fosters cooperative homeostatic loops that improve predictive capabilities and establish functional relationships [98,99].

Chemical reactions can become increasingly complex either through the accumulation of catalytic mutations or by coupling chemical entities of varying complexity. Either pathway allows the emergence of simple autocatalytic processes represented by hierarchical cycles semantically closed (see diagram (4.9)), equivalent to Deacon's [88] autogen. The impredicative nature of any system closed to efficient causation can be interpreted as nested symbolic loops of computation—i.e. self-reference—that are transmuted into non-symbolic processes of self-replication [100]. In this light, self-replication constitutes the threshold for open-ended evolution [101], wherein the evolutionary process itself becomes increasingly unconstrained and generative [102]. Evolution, therefore, emerges as a fundamentally syntax–pragmatic phenomenon, characterized by the unpredictability and emergence of future states. An insight consistent with the trialectic framework articulated by Jaeger *et al.* [68].

A distinction must be drawn between self-replication and self-reproduction. Self-reproduction requires a copier C acting on raw materials [103], whereas in self-replication, the copier C is null, being implicit in the laws of physics [86]. Examples of self-replication include crystals, short RNA strands and autocatalytic cycles. In resource-limited environments, chemical networks compete, self-replicate and mutate through minor variations in composition, ensuring that each type of replicator produces at least one viable offspring, on average, per lifetime. This constitutes a primitive form of evolution [86], characterized by natural selection with limited inheritance due to low robustness and accuracy in replication.

Another important consideration is that, even at this primordial stage, geophysical and chemical processes are naturally cyclical, ensuring a constant energized, non-equilibrium, dynamic exchange between environment and self-replicators. Under these conditions, dynamic kinetic stability (DKS), a kinetic form of stability associated with entities capable of self-replication, emerges [104]. DKS guides self-replication towards greater robustness [105], enabling increasingly complex dynamics to arise as self-replicators proliferate [106], preserving encoded environmental information and preventing its deterioration.

As we pointed out before, the syntactic, semantic and pragmatic dimensions of information are inseparable. Although the presence of semantic (or pragmatic) information in such a primordial state may seem improbable, meaning operates at multiple scales, being syntax–pragmatic interactions its only requirement for existence [107]. According to Küppers [91], the protosemantics of biological information is defined by the carrier's capacity to reproduce rapidly while maintaining high accuracy and stability. This aligns with DKS and probably facilitated the development of sophisticated semantic closure [27].

Over time, the process of infotaxis probably also facilitated compartmentalization, enabling chemical networks to associate mutually, enhancing their functionality and robustness. This progression can be represented by diagrams like (4.4), where formal causes are inherent within material causes [72]. As these networks became more complex, they began to evolve formal causes inherent to efficient causes, such as f_2 in diagram (4.3). However, purely analogue systems exhibited limited robustness and evolutionary potential compared with analogue systems supplemented by digital information control [33].

Digitization emerged as a natural outcome of the evolution of analogue reaction networks, eventually resembling systems like diagram (4.2), where I^* acts as a free-standing formal cause. This transition allowed chemical networks to communicate more efficiently among themselves and with their progeny, being capable of developing sophisticated mechanisms to sense and respond to the environment, as shown in diagram (4.1). However, the emergence of truly open-ended evolution required more than enhanced signalling or compartmentalized inheritance. It demanded the appearance of systems capable of constructing new

functional configurations from internally stored, symbol-like instructions, thus enabling not merely replication, but the continuous rewriting of what it means to persist. Such systems exhibit a form of universality grounded in discrete, modular and syntactically governed structures; codes that are not just molecularly embodied but materially inert, enabling symbolic reuse across vastly different contexts [108].

As these symbolic capacities expanded, the biosphere itself began to resemble a distributed, asynchronous computation over possible adaptations: not in the sense of simulating all paths, but in pruning the unviable ones through irreversible commitments like extinction [101]. In this setting, survival is not merely the preservation of form, but the preservation of a semantic loop. A closure between symbolic description, constructed function and environmental viability. What emerges is not just complexity, but a form of undecidable elaboration: a dynamical process in which the criteria for persistence are themselves continuously evolving, grounded in the dual logic of physical constraint and symbolic freedom.

Thus, we observe a continuum of increasing complexity, beginning with simple chemical reactions modelled as finite-state automata (see diagram (4.5)) and culminating in organisms with cellular-like architectures, such as the extended (F, A) system depicted in diagram (4.1), which exhibit nested loops of computation. In this way, our model explains the experimental observations made by Dueñas-Díez & Pérez-Mercader [90]. Digitization would have been a gradual process, taking substantial time to unfold. Ultimately, digital life forms may have become dominant, as they are intrinsically more robust and better equipped to survive in the long term, representing the lasting products of life's evolutionary processes [33].

6. Discussion

Through our theoretical model, we have outlined a sequential and hierarchical pathway for the origin and early evolution of semantic closure in nature. As noted by Jaeger et al. [68], relevance realization is a process that cannot be fully captured algorithmically. Our narrative complements this view by situating autopoiesis and adaptation within a relational model that is closed semantically while remaining open syntactically. A simple temporal unfolding of this model equips even the simplest life forms with anticipatory mechanisms, validating autopoiesis, adaptation and anticipation based on active inference as syntax–pragmatic mutualistic processes [16].

While Jaeger et al. [68] highlight the apparent non-computability of relevance realization, their conception of computation centres around von Neumann architecture and Turing machines, ignoring computing architectures capable of implementing measurement and control in their operation, which could reveal learning processes mirroring those observed in living systems. We think this narrative is more meaningful through the lens of computational enactivism [89]. This perspective anchors the active inference mechanisms of organisms in the temporal parametrization of our relational model, taking into account both syntactic (computational) and pragmatic (enactive) aspects of life.

Our model successfully captures numerous properties characteristic of living systems, with the notable exception of controlled growth [45,109]. However, a recent study proposes a model inspired by Rosen's original view of (M, R) systems, capturing life as bipartite directed multigraphs [110]. This approach allows modelling controlled growth. As future work, one idea might be to relate their perspective to ours. Crucially, our model is flexible and independent of specific conditions, being agnostic to the substrate, period or location in which life might have emerged [111–114]. The only fundamental environmental prerequisite in our framework is the presence of a primordial language that enables the processes underlying life to take root. This perspective aligns with the idea that symbol grounding is a necessary precursor to interpretation, as emphasized by Pattee [87].

As Cárdenas et al. [115] demonstrated, the absence of strict hierarchies in organisms enables closure to efficient causation in relational models, a principle upheld in our extended (F, A) system. This makes circular causality a cornerstone of our narrative, which is fully compatible with the principle of biological relativity [116,117]. Biological relativity posits that no single level of causation is privileged to describe life, ruling out phenomena such as downward causation. Instead, this aligns with the concept of causal spreading, which suggests that any component within a biological system can influence any other, depending on context [118].

While our narrative illustrates how primitive mechanisms of information control in life might have evolved, it does not claim to pinpoint the precise moment life began. Depending on the definition we use [119], life could be said to arise at various points in our framework. From a gradualist perspective, this ambiguity is not problematic [91,120]. However, it is clear that after the emergence of structures closed to efficient causation (and, as a consequence, semantically closed), many of the defining properties of life emerged. This marks what we might call a self-referential takeover, as chemical systems beyond this point acquire nested loops of computation. Our analysis demonstrates that self-referentiality is not only a necessary condition for life but also essential even for the simplest autocatalytic systems.

7. Conclusion

In this paper, we extended the (F, A) systems proposed by Hofmeyr [71] through a temporal parametrization. This extension captures not only fundamental properties of life, such as thermodynamic openness, catalytic specificity, catalytic and structural closure, information handling, reproduction and evolvability, but also anticipation. Using this relational biological model, we systematically derived weaker versions that reveal a hierarchy of processes by which chemical systems acquired increasingly advanced computational (syntactic) and enactive (pragmatic) capabilities, consistent with the experimental findings of Dueñas-Díez & Pérez-Mercader [90].

Semantic closure emerges as a key factor for robustness, self-replication and, thus, the open-ended evolutionary dynamics observed in living systems. Furthermore, by attributing a self-referential nature to life, our computational enactivist framework

aligns with Jaeger *et al.*'s [68] resolution of the problem of relevance. While life (and its consequences, such as agency and cognition) cannot be represented by von Neumann architectures or Turing machines, our work encourages the exploration of architectures that explicitly implement measurement, control and stochasticity to realize life.

Some problems remain open for future exploration. First, the formulation of a Markov blanket for the asynchronous dynamic Bayesian network associated with our temporal parametrization could provide critical insights into relationships among emerging theoretical frameworks aimed at constructing a general theory of agency [82], providing a concrete model for the sensorimotor autonomy described by Kiverstein *et al.* [81]. Second, similarly to what we did with our extended (F, A) systems, we can derive the temporal parametrizations and joint distributions for weak versions of our model, glimpsing the emergence of internal models that guide adaptive action not just in (M, R) systems, but across the entire zoo of systems closed to efficient causation discussed by Hofmeyr [71]. Third, once a taxonomy of semantically closed systems has been provided, we could examine the consequences of having two or more relational entities occupying the same environment, forming an ecology where multiple cellular-like systems can be coupled to build multicellular organisms.

Our work represents a step towards the unification of three intellectual traditions: relational biology, physical biosemiotics and ecological psychology, each of which fundamentally challenges mainstream reductionist and mechanistic views of life and cognition. Intriguingly, all three paradigms originated within the intellectual landscape of the northeastern United States, reflecting a shared historical and geographical milieu that fostered a relational, systemic and meaning-centred conception of living systems. By situating our extended (F, A) systems at the intersection of these traditions, we have articulated a cohesive theoretical synthesis, tentatively termed 'The Northeastern Synthesis', highlighting the profound relevance of relational autonomy, semiotic processes and affordance realization in biological information handling. In line with Stewart's [121] constructivist perspective, we propose that our narrative has the potential to explain a wide range of phenomena, from the origin of life to high-level cognition.

References

1. Pearson K. 1892 *The grammar of science*. London, UK: W. Scott.

2. Schrodinger E. 1946 *What is life? The physical aspect of the living cell*. New York, NY, USA: The Macmillan company.

3. Pattee HH. 1969 Physical problems of heredity and evolution. *Towards Theor. Biol.* **2**, 268–284.

4. Kendrew JC. 1967 Review of *Phage and the origins of molecular biology*. *Sci. Am.* **216**, 142.

5. Bar-Yam Y. 1997 *Dynamics of complex systems*. In *Studies in nonlinearity*. Boulder, CO, USA: Westview Press.

6. Mitchell M. 2009 *Complexity: a guided tour*. Oxford, UK: Oxford University Press.

7. De Domenico M *et al*. 2019 Complexity explained: a grassroot collaborative initiative to create a set of essential concepts of complex systems. See https://complexityexplained.github.io.

8. Ladyman J, Wiesner K. 2020 What is a complex system? In *What is a complex system?*, pp. 117–134. New Haven, CT, USA: Yale University Press. (doi:10.12987/yale/9780300251104.003.0005)

9. Strogatz SH. 2024 *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Boca Raton, FL, USA: Chapman, Hall/CRC.

10. Walters P. 2000 *An introduction to ergodic theory*. vol. 79. New York, NY, USA: Springer Science & Business Media.

11. Bak P, Tang C, Wiesenfeld K. 1988 Self-organized criticality. *Phys. Rev. A* **38**, 364–374. (doi:10.1103/PhysRevA.38.364)

12. Railsback SF, Grimm V. 2019 *Agent-based and individual-based modeling: a practical introduction*. Princeton, NJ, USA: Princeton University Press.

13. Newman M. 2018 *Networks*. Oxford, UK: Oxford University Press.

14. Fudenberg D, Tirole J. 1991 *Game theory*. Cambridge, MA, USA: MIT Press.

15. Pattee HH. 2019 Simulations, realizations, and theories of life. In *Artificial life*, pp. 63–77. New York, NY, USA: Routledge.

16. Pattee HH. The necessity of biosemiotics: matter-symbol complementarity. In *Introduction to biosemiotics: the new biological synthesis* (ed. M Barbieri), pp. 115–132. Dordrecht, The Netherlands: Springer. (doi:10.1007/1-4020-4814-9_4)

17. Pattee HH. 2001 The physics of symbols: bridging the epistemic cut. *BioSystems* **60**, 5–21. (doi:10.1016/s0303-2647(01)00104-6)

18. Barbieri M. 2003 *The organic codes: an introduction to semantic biology*. Cambridge, UK: Cambridge University Press.

19. Barbieri M. 2018 What is code biology? *BioSystems* **164**, 1–10. (doi:10.1016/j.biosystems.2017.10.005)

20. Pattee HH. 1986 Universal principles of measurement and language functions in evolving systems. In *Complexity, language, and life: mathematical approaches* (eds JL Casti, A Karlqvist), pp. 268–281. Berlin, Germany: Springer.

21. Morris CW. 1946 *Signs, language and behavior*. New York, NY, USA: Prentice-Hall.

22. Cariani PA. 1989 On the design of devices with emergent semantic functions. PhD thesis, State University of New York, Binghamton, NY, USA.

23. von Neumann J. 1955 *Mathematical foundations of quantum mechanics*. Princeton, NJ, USA: Princeton University Press.

24. Turing A. 1936 On the computable numbers with an application to the Entscheidungs. Proc. Lond. Math. Soc. **2**, 230–265.

25. Hoffmeyer J, Emmeche C. 1991 Code-duality and the semiotics of nature. In *On semiotic modeling*, vol. **134**. New York, NY, USA: Mouton de Gruyter.

26. Kull K. 1998 On semiosis, umwelt, and semiosphere. *Semiotica* **120**, 299–310.

27. Pattee HH. 1995 Evolving self-reference: matter, symbols, and semantic closure. Commun. Cogn. Artif. Intell. **12**, 1–2.

28. Rosen R. 1987 On the scope of syntactics in mathematics and science: the machine metaphor. In *Real brains: artificial minds*, pp. 1–23. Amsterdam, The Netherlands: North-Holland.

29. Emmeche C. 1996 *The garden in the machine: the emerging science of artificial life*. Princeton, NJ, USA: Princeton University Press.

30. Grozinger L *et al*. 2019 Pathways to cellular supremacy in biocomputing. *Nat. Commun*. **10**, 5250. (doi:10.1038/s41467-019-13232-z)

31. Calude C, Casti J, Dinneen MJ. 1998 *Unconventional models of computation*. Singapore: Springer Science & Business Media.

32. Pattee HH, Raczaszek Leonardi J. 2012 Cell psychology: an evolutionary approach to the symbol-matter problem. In *Laws, language and life: Howard Pattee's classic papers on the physics of symbols with contemporary commentary*, pp. 165–179. Berlin, Germany: Springer.

33. Walker SI, Davies PC. 2013 The algorithmic origins of life. *J. R. Soc. Interface* **10**, 20120869. (doi:10.1098/rsif.2012.0869)

34. Ashby WR. 1949 The electronic brain. *Radio Electronics* 77–79.

35. Rosen R. 1958 A relational theory of biological systems. *Bull. Math. Biophys*. **20**, 245–260. (doi:10.1007/BF02478302)

36. Rosen R. 1958 The representation of biological systems from the standpoint of the theory of categories. *Bull. Math. Biophys*. **20**, 317–341. (doi:10.1007/BF02477890)

37. von Neumann J. 1966 *Theory of self-reproducing automata*. (ed. AW Burks). Urbana, IL, USA: University of Illinois Press.

38. Varela FJ, Maturana HR, Uribe R. 1974 Autopoiesis: the organization of living systems, its characterization and a model. *BioSystems* **5**, 187–196.

39. Maturana H, Varela F. 1980 *Autopoiesis and cognition: the realization of living*. Dordrecht, The Netherlands: Reidel Publishing Company.

40. Prigogine I. 1978 Time, structure, and fluctuations. *Science* **201**, 777–785. (doi:10.1126/science.201.4358.777)

41. Kauffman SA. 1986 Autocatalytic sets of proteins. *J. Theor. Biol*. **119**, 1–24. (doi:10.1016/S0022-5193(86)80047-9)

42. Maley SP, Gershenson C, Kauffman SA. 2025 Life finds a way: emergence of cooperative structures in adaptive threshold networks. *SSRN*. (doi:10.2139/ssrn.5525612)

43. Gánti T. 2003 *Chemoton theory: theory of living systems*. New York, NY, USA: Springer Science & Business Media.

44. EigenM S. 2012 *The hypercycle: a principle of natural self-organization*. Berlin, Germany: Springer Science & Business Media.

45. Cornish-Bowden A, Cárdenas ML. 2020 Contrasting theories of life: historical context, current theories. In search of an ideal theory. *BioSystems* **188**, 104063. (doi:10.1016/j.biosystems.2019.104063)

46. Korbak T. 2023 Self-organisation, (M, R)–systems and enactive cognitive science. *Adapt. Behav*. **31**, 35–49. (doi:10.1177/10597123211066155)

47. Varenne F. 2013 The mathematical theory of categories in biology and the concept of natural equivalence in Robert Rosen. *Revue d'histoire des Sciences* **66**, 167–197. (doi:10.3917/rhs.661.0167)

48. Rosen R. 1991 *Life itself: a comprehensive inquiry into the nature, origin, and fabrication of life*. New York, NY, USA: Columbia University Press.

49. Rashevsky N. 1954 Topology and life: in search of general mathematical principles in biology and sociology. *Bull. Math. Biophys*. **16**, 317–348. (doi:10.1007/BF02484495)

50. Baianul I, Marinescu M. 1974 On a functorial construction of (M, R)-systems. Rev. Roum. Math. Pures. Appl. **19**, 388–391.

51. Eilenberg S, MacLane S. 1945 General theory of natural equivalences. *Trans. Am. Math. Soc*. **58**, 231–294. (doi:10.1090/s0002-9947-1945-0013131-6)

52. Rosen R. 1971 Some realizations of (M, R)-systems and their interpretation. *Bull. Math. Biophys*. **33**, 303–319. (doi:10.1007/BF02476776)

53. Rosen R. 1978 On anticipatory systems: I. When can a system contain a predictive model of another? J. Soc. Biol. Syst. **1**, 155–162. (doi:10.1016/S0140-1750(78)80003-7)

54. Letelier JC, Cárdenas ML, Cornish-Bowden A. 2011 From *L'Homme Machine* to metabolic closure: steps towards understanding life. *J. Theor. Biol*. **286**, 100–113. (doi:10.1016/j.jtbi.2011.06.033)

55. Cárdenas ML, Letelier JC, Gutierrez C, Cornish-Bowden A, Soto-Andrade J. 2010 Closure to efficient causation, computability and artificial life. *J. Theor. Biol*. **263**, 79–92. (doi:10.1016/j.jtbi.2009.11.010)

56. Letelier JC *et al*. 2010 (M, R) systems and RAF sets: common ideas, tools and projections. In *Proc. of the Alife XII Conf*, Odense, Denmark.

57. Letelier JC, Soto-Andrade J, Guíñez Abarzúa F, Cornish-Bowden A, Luz Cárdenas M. 2006 Organizational invariance and metabolic closure: analysis in terms of (M, R) systems. *J. Theor. Biol*. **238**, 949–961. (doi:10.1016/j.jtbi.2005.07.007)

58. Lane PA. 2024 Robert Rosen's relational biology theory and his emphasis on non-algorithmic approaches to living systems. *Mathematics* **12**, 3529. (doi:10.3390/math12223529)

59. Pattee HH. 2007 Laws, constraints, and the modeling relation – history and interpretations. *Chem. Biodivers*. **4**, 2272–2295. (doi:10.1002/cbdv.200790187)

60. Rosen R. 1964 Abstract biological systems as sequential machines. *Bull. Math. Biophys*. **26**, 103–111. (doi:10.1007/BF02476831)

61. Rosen R. 1964 Abstract biological systems as sequential machines. II. Strong connectedness and reversibility. *Bull. Math. Biophys*. **26**, 239–246. (doi:10.1007/BF02479045)

62. Rosen R. 1966 Abstract biological systems as sequential machines. III. Some algebraic aspects. *Bull. Math. Biophys*. **28**, 141–148. (doi:10.1007/BF02476987)

63. Soto-Andrade J, Jaramillo S, Gutiérrez C, Letelier JC. 2011 Ouroboros avatars: a mathematical exploration of self-reference and metabolic closure. In *Advances in Artificial Life ECAL 2011: Proc. of the 11th European Conf. on the Synthesis and Simulation of Living Systems*, pp. 763–770. Cambridge, MA, USA: The MIT Press.

64. Mossio M, Longo G, Stewart J. 2009 A computable expression of closure to efficient causation. *J. Theor. Biol*. **257**, 489–498. (doi:10.1016/j.jtbi.2008.12.012)

65. Palmer ML, Williams RA, Gatherer D. 2016 Rosen's (M, R) system as an X-machine. *J. Theor. Biol*. **408**, 97–104. (doi:10.1016/j.jtbi.2016.08.007)

66. Gatherer D, Galpin V. 2013 Rosen's (M, R) system in process algebra. *BMC Syst. Biol*. **7**, 1–11. (doi:10.1186/1752-0509-7-128)

67. Zhang L, Williams RA, Gatherer D. 2016 Rosen's (M, R) system in unified modelling language. *BioSystems* **139**, 29–36. (doi:10.1016/j.biosystems.2015.12.006)

68. Jaeger J, Riedl A, Djedovic A, Vervaeke J, Walsh D. 2024 Naturalizing relevance realization: why agency and cognition are fundamentally not computational. *Front. Psychol*. **15**, 1362658. (doi:10.3389/fpsyg.2024.1362658)

69. Rosen R. 1963 Some results in graph theory and their application to abstract relational biology. *Bull. Math. Biophys*. **25**, 231–241. (doi:10.1007/BF02478280)

70. Horowitz NH. 1948 The one gene-one enzyme hypothesis. *Genetics* **33**, 612.

71. Hofmeyr JHS. 2021 A biochemically-realisable relational model of the self-manufacturing cell. *BioSystems* **207**, 104463. (doi:10.1016/j.biosystems.2021.104463)

72. Hofmeyr JHS. 2018 Causation, constructors and codes. *BioSystems* **164**, 121–127. (doi:10.1016/j.biosystems.2017.09.008)

73. Kauffman S, Roli A. 2021 The world is not a theorem. *Entropy* **23**, 1467. (doi:10.3390/e23111467)

74. Pattee HH, Raczaszek-Leonardi J, Pattee HH. 2012 Clues from molecular symbol systems. In *Laws, language and life: Howard Pattee's classic papers on the physics of symbols with contemporary commentary*, pp. 155–163. Berlin, Germany: Springer.

75. Barbieri M. 1981 The ribotype theory on the origin of life. *J. Theor. Biol*. **91**, 545–601. (doi:10.1016/0022-5193(81)90211-3)

76. Gibson JJ. 1979 *The ecological approach to visual perception*. Boston, MA, USA: Houghton Mifflin.

77. Montévil M, Mossio M. 2015 Biological organisation as closure of constraints. *J. Theor. Biol*. **372**, 179–191. (doi:10.1016/j.jtbi.2015.02.029)

78. Friston K. 2013 Life as we know it. *J. R. Soc. Interface* **10**, 20130475. (doi:10.1098/rsif.2013.0475)

79. Clark A. 2017 How to knit your own markov blanket: resisting the second law with metamorphic minds. In *Philosophy and predictive processing* (eds TK Metzinger, W Wiese), p. 3. Frankfurt am Main, Germany: MIND Group.

80. Biehl M, Virgo N. 2022 Interpreting systems as solving pomdps: a step towards a formal understanding of agency. In *International Workshop on Active Inference*, pp. 16–31. Cham, Switzerland: Springer.

81. Kiverstein J, Kirchhoff MD, Froese T. 2022 The problem of meaning: the free energy principle and artificial agency. *Front. Neurorobotics* **16**, 844773. (doi:10.3389/fnbot.2022.844773)

82. Egbert M, Hanczyc MM, Harvey I, Virgo N, Parke EC, Froese T, Sayama H, Penn AS, Bartlett S. 2023 Behaviour and the origin of organisms. In *Origins of life and evolution of biospheres* (eds M Egbert, MM Hanczyc, I Harvey), pp. 1–26. Dordrecht, The Netherlands: Springer Science and Business Media B.V.

83. Barbieri M. 2019 Code biology, Peircean biosemiotics, and Rosen's relational biology. *Biol. Theory* **14**, 21–29. (doi:10.1007/s13752-018-0312-z)

84. Pattee HH. 1969 How does a molecule become a message? In *Communication in development* (ed. CH Waddington), pp. 1–16. Cambridge, MA: Academic Press. (doi:10.1016/B978-0-12-395541-8.50006-9)

85. Waters DP. 2012 Von Neumann's theory of self-reproducing automata: a useful framework for biosemiotics? *Biosemiotics* **5**, 5–15. (doi:10.1007/s12304-011-9127-z)

86. Marletto C. 2015 Constructor theory of life. *J. R. Soc. Interface* **12**, 20141226. (doi:10.1098/rsif.2014.1226)

87. Pattee HH. 2021 Symbol grounding precedes interpretation. *Biosemiotics* **14**, 561–568. (doi:10.1007/s12304-021-09458-4)

88. Deacon TW. 2021 How molecules became signs. *Biosemiotics* **14**, 537–559. (doi:10.1007/s12304-021-09453-9)

89. Korbak T. 2021 Computational enactivism under the free energy principle. *Synthese* **198**, 2743–2763. (doi:10.1007/s11229-019-02243-4)

90. Dueñas-Díez M, Pérez-Mercader J. 2019 How chemistry computes: language recognition by non-biochemical chemical automata. From finite automata to Turing machines. *iScience* **19**, 514–526. (doi:10.1016/j.isci.2019.08.007)

91. Küppers BO. 1990 *Information and the origin of life*. Cambridge, MA, USA: MIT Press.

92. Segré D, Lancet D. 2000 Composing life. *EMBO Rep.* **1**, 217–222. (doi:10.1093/embo-reports/kvd063)

93. Hunding A, Kepes F, Lancet D, Minsky A, Norris V, Raine D, Sriram K, Root-Bernstein R. 2006 Compositional complementarity and prebiotic ecology in the origin of life. *Bioessays* **28**, 399–412. (doi:10.1002/bies.20389)

94. Markovitch O, Lancet D. 2012 Excess mutual catalysis is required for effective evolvability. *Artif. Life* **18**, 243–266. (doi:10.1162/artl_a_00064)

95. Hopcroft JE, Motwani R, Ullman JD. 2001 Introduction to automata theory, languages, and computation. *SIGACT News* **32**, 60–65. (doi:10.1145/568438.568455)

96. Bartlett S, Louapre D. 2022 Provenance of life: chemical autonomous agents surviving through associative learning. *Phys. Rev. E* **106**, 034401. (doi:10.1103/physreve.106.034401)

97. Poole W, Ortiz-Munoz A, Behera A, Jones NS, Ouldridge TE, Winfree E, Gopalkrishnan M. 2017 Chemical Boltzmann machines. In *DNA Computing and Molecular Programming: 23rd Int. Conf., DNA 23*, Austin, TX, USA, pp. 210–231. Cham, Switzerland: Springer. (doi:10.1007/978-3-319-66799-7_14)

98. Levin M. 2019 The computational boundary of a 'self': developmental bioelectricity drives multicellularity and scale-free cognition. *Front. Psychol.* **10**, 493866. (doi:10.3389/fpsyg.2019.02688)

99. Heins C, Millidge B, Da Costa L, Mann RP, Friston KJ, Couzin ID. 2024 Collective behavior from surprise minimization. *Proc. Natl. Acad. Sci.* **121**, e2320239121. (doi:10.1073/pnas.2320239121)

100. Sayama H. 2008 Construction theory, self-replication, and the halting problem. *Complexity* **13**, 16–22. (doi:10.1002/cplx.20218)

101. Hernández-Orozco S, Hernández-Quiroz F, Zenil H. 2018 Undecidability and irreducibility conditions for open-ended evolution and emergence. *Artif. Life* **24**, 56–70. (doi:10.1162/artl_a_00254)

102. Pattee HH, Sayama H. 2019 Evolved open-endedness, not open-ended evolution. *Artif. Life* **25**, 4–8. (doi:10.1162/artl_a_00276)

103. Szathmáry E. 1995 A classification of replicators and lambda-calculus models of biological organization. *Proc. R. Soc. Lond. B* **260**, 279–286. (doi:10.1098/rspb.1995.0092)

104. Pascal R, Pross A. 2015 Stability and its manifestation in the chemical and biological worlds. *Chem. Commun.* **51**, 16160–16165. (doi:10.1039/C5CC06260H)

105. Pascal R, Pross A, Corning P, Kauffman S, Noble D, Shapiro J, Vane-Wright R, Pross A. 2023 Toward the physicalization of biology: seeking the chemical origin of cognition. In *Evolution 'on purpose': teleonomy in living systems* (eds PA Corning, SA Kauffman, D Noble, JA Shapiro, RI Vane-Wright, A Pross), pp. 257–274. Cambridge, MA, USA: Vienna Series Theoretical Biology. (doi:10.7551/mitpress/14642.003.0017)

106. Alakuijala J, Evans J, Laurie B, Mordvintsev A, Niklasson E, Randazzo E, Versari L. 2024 Computational life: how well-formed, self-replicating programs emerge from simple interaction. *arXiv* (doi:10.48550/arXiv.2406.19108)

107. Fields C, Levin M. 2020 How do living systems create meaning? *Philosophies* **5**, 36. (doi:10.3390/philosophies5040036)

108. Rocha LM, Hordijk W. 2005 Material representations: from the genetic code to the evolution of cellular automata. *Artif. Life* **11**, 189–214. (doi:10.1162/1064546053278964)

109. Cornish-Bowden A, Cárdenas ML. 2022 The essence of life revisited: how theories can shed light on it. *Theory Biosci.* **141**, 105–123. (doi:10.1007/s12064-021-00342-w)

110. Márquez-Zacarías P, Ortiz-Muñoz A, Bingham EP. 2025 The nature of organization in living systems. *ArXiv* 2503.03950v1. (doi:10.48550/arXiv.2503.03950)

111. Abbot DS, Switzer ER. 2011 The Steppenwolf: a proposal for a habitable planet in interstellar space. Astrophys. J. Lett. **735**, L27. (doi:10.1088/2041-8205/735/2/L27)

112. Loeb A. 2014 The habitable epoch of the early universe. *Int. J. Astrobiology* **13**, 337–339. (doi:10.1017/S1473550414000196)

113. Pearce BKD, Tupper AS, Pudritz RE, Higgs PG. 2018 Constraining the time interval for the origin of life on Earth. *Astrobiology* **18**, 343–364. (doi:10.1089/ast.2017.1674)

114. Ballesteros FJ, Fernandez-Soto A, Martínez VJ. 2019 Diving into exoplanets: are water seas the most common? *Astrobiology* **19**, 642–654. (doi:10.1089/ast.2017.1720)

115. Cárdenas ML, Benomar S, Cornish-Bowden A. 2018 Rosennean complexity and its relevance to ecology. *Ecol. Complex.* **35**, 13–24. (doi:10.1016/j.ecocom.2017.04.005)

116. Noble D. 2012 A theory of biological relativity: no privileged level of causation. *Interface Focus* **2**, 55–64. (doi:10.1098/rsfs.2011.0067)

117. Noble R, Tasaki K, Noble PJ, Noble D. 2019 Biological relativity requires circular causality but not symmetry of causation: so, where, what and when are the boundaries? *Front. Physiol.* **10**, 827. (doi:10.3389/fphys.2019.00827)

118. Ball P. 2023 What distinguishes the elephant from *E. coli*: causal spreading and the biological principles of metazoan complexity. *J. Biosci.* **48**, 14. (doi:10.1007/s12038-023-00344-4)

119. Benner SA. 2010 Defining life. *Astrobiology* **10**, 1021–1030. (doi:10.1089/ast.2010.0524)

120. Gershenson C. 2012 The world as evolving information (eds AA Minai, D Braha, Y Bar-Yam). In *Unifying Themes in Complex Systems VII: Proc. of the 7th Int. Conf. on Complex Systems*, pp. 100–115. Berlin, Germany: Springer.

121. Stewart J. 1995 Cognition = life: implications for higher-level cognition. *Behav. Process.* **35**, 311–326. (doi:10.1016/0376-6357(95)00046-1)