Regulative development as a model for origin of life and artificial life studies

Chris Fields[a*] and Michael Levin[a,b]

[a] *Allen Discovery Center at Tufts University, Medford, MA 02155 USA*
ORCID: 0000-0002-4812-0744
[b] *Wyss Institute for Biologically Inspired Engineering at Harvard University,*
Boston, MA 02115, USA
ORCID: 0000-0001-7292-8084

May 21, 2023

Abstract

Using the formal framework of the Free Energy Principle, we show how generic thermodynamic requirements on bidirectional information exchange between a system and its environment can generate complexity. This leads to the emergence of hierarchical computational architectures in systems that operate sufficiently far from thermal equilibrium. In this setting, the environment of any system increases its ability to predict system behavior by "engineering" the system towards increased morphological complexity and hence larger-scale, more macroscopic behaviors. When seen in this light, regulative development becomes an environmentally-driven process in which "parts" are assembled to produce a system with predictable behavior. We suggest on this basis that life is thermodynamically favorable and that, when designing artificial living systems, human engineers are acting like a generic "environment".

---

*Corresponding author at: Allen Discovery Center at Tufts University, Medford, MA 02155 USA; *E-mail address*: fieldsres@gmail.com

1 Introduction

Living systems are hierarchical arrangements of components that, critically, exhibit a multi-scale competency architecture (MCA; Levin, 2022; Clawson and Levin, 2022). In an MCA, components at each scale are competent to perform the functions appropriate to that scale without explicit top-down instructions. Human cells, for example, do not have to be told how to divide by the brain, or by any larger-scale system; they are competent to divide on their own. At any scale j in an MCA, the behaviors of the components at the lower scale $j - 1$ provide enabling mechanisms for j-appropriate behaviors, while the behaviors of the components at the higher scale $j + 1$ impose boundary conditions. As the boundary conditions effectively constrain the possibility space at scale j, state changes at scale $j + 1$ can modulate or deform the possibility space at scale j, providing a source of top-down novelty. Cell division, for example, is enabled by gene expression within the cell, and constrained by signaling from, or transmitted via, the surrounding multicellular tissue. Top-down "instructions" from the surrounding, or even distal (McMillan et al., 2021), tissue can increase or decrease the rate of cell division, e.g. during wound healing; other examples include the modulation of cancer cell properties by the microenvironment (Bissell et al., 2002; Ingber, 2008; Bizzarri and Cucina, 2014), control of stem cell fate by large-scale axial patterning cues during regeneration (Durant et al., 2017), and local remodeling of a structure based on its larger anatomical context (Farinella-Ferruzza, 1956).

Such upward flows of enabling mechanisms and downward flows of constraining boundary conditions have been identified as characteristic of living systems by Polanyi (1968), Rosen (1986), Pezzulo and Levin (2016), and by us (Fields and Levin, 2020a) among others. In the language of cognitive or computer science, an MCA encapsulates the competencies required for scale-appropriate behaviors. Such scale-specific encoding avoids, in particular, "micromanagement" in the form of explicit top-down directives for each step of a complex, lower-level competency. The automatization of processes such as first-language understanding and production – and of many skills that initially require explicit rule learning (Bargh and Ferguson, 2000) – provides a familiar example. Scale-specific encapsulation of competencies allows processes specific to each scale to function as virtual machines (Smith and Nair, 2005), independently of the implementations of either lower- or higher-scale components. Such implementation-independence opens up the possibility of "mix and match" systems that combine evolved biological, engineered biological, and artificial components in almost arbitrary ways (Levin, 2022; Clawson and Levin, 2022).

Making the idea of an MCA precise requires having a precise formulation of what counts as a "competence." To develop a fully-precise formulation of competence, we turn to the variational Free Energy Principle (FEP). The FEP was first introduced as a theory of brain function (Friston, 2005; Friston et al., 2006; Friston, 2010), was subsequently developed into a theoretical framework for modeling living systems (Friston, 2013; Friston et al., 2017; Ramstead et al., 2019; Kuchling et al., 2020), and more recently, extended into a framework for modeling any physical system that remains distinguishable from its environment over some time period of interest (Friston, 2019; Ramstead et al., 2022). These formulations

employ the classical theory of random dynamical systems; expressed in this way, the FEP is a classical least-action principle. In this classical setting, distinguishability between the system of interest and its environment is assured by requiring that the two be separated by a Markov blanket (MB; Pearl, 1988; Clark, 2017), as discussed in §2.1 below. Using quantum information-theoretic methods, the MB condition can be further generalized to the condition of separability, the requirement that the joint system-environment state can be factored into independently-specifiable system and environment states. Separability is both necessary and sufficient for an effectively classical interaction between the system and its environment (Fields and Glazebrook, 2020; Fields and Marcianò, 2020; Fields, Glazebrook and Marcianò, 2021). In this quantum-information formulation, minimizing VFE optimizes the interface between two computational processes, one implemented by the system and the other by its environment (Fields et al., 2022). In addition, as discussed in §2.2 below, the FEP emerges as the classical limit of the Principle of Unitary, i.e. the Principle of Conservation of Information.

In the context of the FEP, a system is "competent" if and only if it is able to maintain its distinguishability from its environment over the time period of interest by maintaining the functional integrity of its MB (or in quantum terms, its separability from its environment). As discussed in §2.1, a system's competency is measured by its ability to minimize a variational free energy (VFE) that is defined at its MB. In this case, VFE is effectively an uncertainty about what its environment will do next. In other words, highly competent systems are highly competent at both predicting what their environments will do next, and responding to their environments' behaviors in a way that preserves their predictive power. Friston (2019) refers to such systems as both self-organizing and "self-evidencing," i.e. as continually generating evidence of their own continuing existence.

By providing a scale-free, effectively thermodynamic definition of competence, the FEP enables a scale-free biology that treats evolution and development as manifestations (albeit at different levels of organization) of a single process – VFE minimization (Fields and Levin, 2020a,b). It is natural to extend this conception of biology to encompass biologically-relevant prebiotic, abiotic, or exobiotic processes, including those invoked in origin of life and artificial life scenarios. To investigate the definition of "competence" provided by the FEP in this larger context, we will focus, in particular, on two kinds of biological self-organization:

> **Regulative development:** Using information available from their surrounding environment(s), some number of cells self-organize into a multicellular organism. In both embryonic development and regeneration, such regulated self-organization enables multicellular systems to reach "normal" target morphologies despite significant perturbations or altered starting conditions (Birnbaum and Alvarado, 2008; Levin, 2011; Vandenberg, Adams and Levin, 2012; Lobo et al., 2014; Pezzulo and Levin, 2016; Pinet and McLaughlin, 2019; Fields and Levin, 2020b). Such processes are also employed by heterogeneous, facultative

multicellular systems like microbial mats.

Ab initio self organization: Some number of molecules self-organize into a cell. The result could also be a self-sustaining proto-cell as in some origin-of-life models.

Regulative development has been studied for over two centuries, and many model systems are now understood at a substantial level of detail including, in several cases, cell-specific gene expression profiles (Tintori et al., 2016; Farnsworth, Saunders and Miller, 2020; Wang et al., 2022). Examples of *ab initio* self-organization, on the other hand, remain hypothetical. While a variety of models have been proposed, mostly in an origin-of-life context, none have been fully implemented experimentally and there is little consensus even about biological plausibility. The goal of any scale-free explanation is to tell exactly the same kind of story about these two examples. Constructing such an explanation by requiring examples of *ab initio* self-organization to be both formal and mechanistic analogs of regulative development raises a number of issues, mainly concerning the structure and role of the "environment" in *ab initio* self-organization, that have been relatively neglected by previous approaches.

In particular, we will see that treating *ab initio* self-organization as an analog of regulative development challenges two deeply-entrenched ideas. First, it questions the near-universal assumption that any *ab initio* model must result in a self-replicating system. Evolutionary models from Darwin onwards have strongly coupled variation with inheritance via either meiotic or mitotic cell division (Monod, 1972; Szathmáry and Maynard Smith, 1995; Michod, 1999). Natural selection acts on variants, and hence depends on Darwinian models of reproduction. Such models suggest *ab initio* processes that involve self-replicating molecules immediately. We will suggest that models in which the needed "parts" are generated by environmental processes that are at most weakly coupled to the systems of interest are also worth consideration. Variation generated by weakly coupled processes is consistent with evolution at the global scale, but does not depend on natural selection at any single local scale; it thus provides a "Non-Darwinian" source of order. Such weak coupling is exemplified by situations involving self-organizing systems that include engineered and manufactured components; we suggest that weak-coupling models may be realistic in other settings as well.

Second, treating *ab initio* self-organization as an analog of regulative development challenges the very idea of *self*-organization. In its purest form, the idea of self-organization suggests that the information needed for organization is present in the self. This immediately raises the essentially unanswerable question of how this critical information got there to begin with. Directly relating *ab initio* self-organization to regulative development forces us to ask, at each step of the process, what systems count as "selves," how the environment of each "self" is defined, and how, in each case, the exchange of information between "self" and environment is implemented (Levin, 2021).

Our analysis suggests two broad conclusions:

1. Reproduction via cell division is an efficiency hack – a cheap heuristic – which evolution "froze in" by selecting strong self/other recognition systems starting in bacteria. Decoupling replication from self-organization generalizes the systems of interest to developmental studies toward "egalitarian" (Strassmann and Queller, 2010) assemblages of unrelated, or in the evolved case, only distantly related, components that "just happen" to work well together. This allows simpler origin of life stories. It acknowledges the possibility of kinematic replication in biological systems, as observed experimentally in "xenobots" constructed from dissociated *Xenopus laevis* skin cells (Kriegman et al., 2021). It suggests that symbiotic systems such as microbial mats are "canonical" to at least the same same extent as obligately sexual multicellulars. The focus of both *ab initio* and developmental studies becomes, in this case, how components interact once they are placed in mutual proximity, independently of their origins.

2. "Self-organization" is always environment-dependent, so we can view it as at least in part environment-directed: some of the instructive information is initially in the environment, not in the self that assembles. From the perspective of physical interaction (Fields and Marcianò, 2020; Fields, Glazebrook and Marcianò, 2021) or of the FEP (Friston, 2019; Ramstead et al., 2022; Fields et al., 2022), this is obvious. It is also obvious in many origin of life models, in microbial mat assembly, and in embryogenesis, even in "European plan" organisms like *C. elegans* (Barrière and Bertrand, 2020). It is, nonetheless, often neglected or de-emphasized, particularly in theoretical work. As discussed in Fields et al. (2022), the FEP applies to the environment of any system of interest, rendering it an uncertainty-minimizing agent as discussed in §4 below. The environment-dependence of self-organization is a consequence of the environment acting as a Bayesian agent.

In what follows, we first review in §2 both the classical development of the FEP and its quantum information-theoretic generalization. As the latter provides the most convenient formalism for describing the system-environment interaction, we adopt it in what follows. We then discuss in §3 a fundamental symmetry of the FEP that is often neglected: the FEP requires the environment of any system to also be a VFE-minimizing agent. While any system-environment interaction is informationally symmetric – equal quantities of information flow in both directions – what the two parties do with the information they receive may be radically different. Symmetric information flows, in other words, are consistent with "cognitive light cones" (Levin, 2021, 2022) of different widths and depths, and hence different active inference capabilities, on the two sides of the system-environment boundary. We turn in §4 to an explicit comparison between regulative development and *ab initio* self-organization, focusing first on characterizing the environments of each active component as information sources, and hence as themselves active agents, at each step in the process. We investigate, in particular, how both regulative development and *ab initio* self-organization decrease the VFE detected by the environment, making them thermodynamically-driven processes under the FEP. The environment of any system, in other words, can be expected to act so as to increase that system's complexity, with the

likely outcome being a complexity-increasing "arms race" between the system and its environment if both have sufficient computational resources. We conclude by outlining a "mix and match" experimental strategy suggested by these results in §5.

2 Classical and quantum formulations of the FEP

2.1 Classical formulations

The FEP applies to physical systems – what Friston (2019) calls "things" or "particles" – that are persistent, in the sense of having a well-defined state space, over some time period of interest. Adopting the notation of Friston (2019), if the internal states μ of such a system S are conditionally independent of the states η of its environment E, it (by definition) possesses an MB, defined as the set b of states on which the dependence of μ on η is conditioned. This MB condition will hold, in general, provided the interaction between system and environment is significantly weaker than the internal self-interactions of either. Three remarks are in order here. First, the environment E is the *entire* environment of S, or at least the entire environment of interest; in what follows, we will always consider E to comprise "everything but S." Second, fixing the sets of states μ and η uniquely fixes the set b and hence the MB. Every system, therefore, has a unique MB separating it from its environment. Third, the definition of an MB makes no explicit reference to ordinary, three-dimensional (3d) space. Therefore, while it is commonplace to consider the MB of an organism, for example, to coincide with or be implemented by its 3d spatial boundary, this is not a requirement of the theory. Indeed, nothing in principle prevents a collection of spatially-disconnected entities, e.g. a population of organisms, from having an MB.

Given the conditions above, the VFE is a statistical relations between internal, external, and intervening blanket states. It can be written Friston (2019), Eq. 2.3:

$$
\begin{aligned}
F(\pi) &= \underbrace{\mathbb{E}_{q(\eta)}[\ln q_\mu(\eta) - \ln p(\eta, b)]}_{\text{Variational free energy}} \\
&= \underbrace{\mathbb{E}_q[-\ln p(b|\eta) - \ln p(\eta)]}_{\text{Energy constraint (likelihood \& prior)}} - \underbrace{\mathbb{E}_q[-\ln q_\mu(\eta)]}_{\text{Entropy}} \\
&= \underbrace{D_{KL}[q_\mu(\eta)|p(\eta)]}_{\text{Complexity}} - \underbrace{\mathbb{E}_q[\ln p(b|\eta)]}_{\text{Accuracy}} \\
&= \underbrace{D_{KL}[q_\mu(\eta)||p(\eta|b)]}_{\text{Divergence}} \underbrace{-\ln p(b)}_{\text{Log evidence}} \geq -\ln p(b)
\end{aligned}
\tag{1}
$$

where $\pi(t) = (\mu(t), b(t))$ is the time-dependent "particle" state. The VFE functional $F(\pi)$ is an upper bound on surprisal (a.k.a. self-information) $\mathfrak{I}(\pi) = -\log P(\pi) = -\ln p(b)$ because the Kullback-Leibler divergence term (D_{KL}) is always non-negative. This KL divergence is between the probability density over external states η, given the MB state b, and a

variational density $q_\mu(\eta)$ over external states parameterized by the internal state μ. Note here that the blanket states b are considered components of the "particle" state π; hence the VFE defined by Eq. (1) is the VFE "for" or "experienced by" S. We could, however, also consider a composite state $\rho = (\eta, b)$ and write an analog of Eq. (1) in which the roles of μ and η are exchanged; the resulting $F(\rho)$ would be the VFE "for" or "experienced by" E. This symmetry of the FEP will be explored further in §3 below.

If the joint system $S \sqcup E$, where $\sqcup$ denotes disjoint union, is represented as a random dynamical system, then S having an MB requires that state trajectories that start in μ remain in μ; the states of S cannot, in other words, diverge to outside of the boundary of S. This condition is met if S has some non-equilibrium steady state (NESS) density. If we view the internal state μ as encoding a posterior probability distribution over the external state η, then minimizing VFE is, effectively, minimizing a prediction error under a generative model (GM) encoded by the NESS density. In this interpretation, Eq. (1) may be viewed as defining a "Bayesian mechanics" (Ramstead et al., 2022) and minimization of VFE is a form of inference, termed "active inference" because one way of minimizing VFE is to act on the environment to move it toward an expected state. In this case, an agent S is competent if and only if it is an effective minimizer of its experienced VFE, i.e. if it can prevent its VFE from becoming high enough to drive its states far from its NESS, destroying the integrity of its MB(Friston, 2019).

While a substantial literature now supports the applicability of the FEP to living systems (see e.g. Friston et al. (2020), Smith, Babcock and Friston (2020) for topical reviews), the assumptions of static NESS densities and MBs have also been criticized as unrealistic for living systems (e.g. Raja et al. (2021); Aguilera et al. (2021); Bruineberg et al. (2022); Di Paolo, Thompson and Beer (2022); see also Biehl, Pollock and Kanai (2021) for a critique of ancillary assumptions related to these). Central to several of these critiques is an emphasis on the history-dependence of living systems and the neglect of history, i.e. encoded memory, in specifying system states under the FEP. In part as a response to these criticisms, the FEP has been reformulated as a least-action principle on paths or trajectories in the joint $S \sqcup E$ state space (Friston et al., 2022), an extension that permits a re-interpretation in terms of classical gauge fields (Sakthivadivel, 2022a,b); see Ramstead et al. (2022) for a summary of these developments. These formulations do not require assuming a NESS or a fixed MB; see Sakthivadivel (2022c), in particular, for a discussion of material exchange between a system and its environment.

We have taken an alternative, but as shown in Fields et al. (2022), fully-consistent approach to generalizing the FEP, reformulating it as a quantum information-theoretic principle. This formulation applies to all physcal systems that are not entangled with, and hence can be distinguished from, their environments. It explicitly addresses the questions of how organisms identify specific objects within their environments, and how they write to and read from their memories of specific events.

2.2   Quantum information-theoretic formulation

As Deutsch (1997) points out, information theory is inherently quantum: it is about answers to yes/no questions (Wheeler, 1989). The idea that quantum theory is fundamentally about observation can be traced back to Bohr (1928); see also Bohr (1958). Fuchs (2003) proposed that quantum theory is fundamentally about observation followed by Bayesian inference; this view has since been productively developed by the "QBist" movement (Fuchs, 2010; Fuchs and Schack, 2013; Mermin, 2018). Quantum information theory represents the joint system $S \sqcup E$ by a Hilbert space, which for simplicity we assume to be finite and denote U for "universe" of interest. As we take E to contain everything (of interest) other than S, we can choose it to be large enough that the joint system U can be considered isolated, or closed. The basis vectors of any Hilbert space represent the possible values of the degrees of freedom of the relevant system; hence the basis vectors of U represent all possible values of all degrees of freedom of either S or E. As a measurement of each degree of freedom either yields, or does yield, any particular value at any particular time, we can treat these basis vectors as binary without loss of generality; hence we can represent $U = \otimes_{i=1}^{M} q_i$, where the dimension $M = \dim(U)$ is the number of basis vectors of U, the q_i are quantum bits (qubits), i.e. Hilbert spaces of dimension one, and $\otimes$ is the Hilbert-space tensor product. We can represent the time-evolution of the state of U, in an abstract, parametric time t, as:

$$P_U = e^{-(\imath/\hbar)H_U t} \tag{2}$$

where H_U is the Hamiltonian (total energy) operator on U and $\hbar$ is Planck's constant; see Nielsen and Chuang (2000) (or any other textbook) for a standard introduction to this formalism.

The Hamiltonian operator is additive; hence for any decomposition of U into subsystems S and E, we can write $H_U = H_S + H_E + H_{SE}$, where H_S and H_E and the internal or "self" interactions of S and E, respectively, and H_{SE} represents the interaction between S and E. We now introduce the single assumption that underlies the FEP:

> **Separability:** We assume that the joint state $|U\rangle$ (employing the Dirac notation for states) factors, during the time period of interest, as $|U\rangle = |SE\rangle = |S\rangle|E\rangle$.

If multiple observations are involved, we require that the density $\rho_U = \rho_{SE} = \rho_S \rho_E$. The joint state, or state density, is separable, i.e. factors, if and only if (indeed, by definition) it is not entangled, i.e. the entanglement entropy $\mathcal{S}(SE)$ across the S-E boundary, which we denote by $\mathscr{B}$, is zero. This condition guarantees that the states of S and E can be independently specified, i.e. they are conditionally independent.

The interaction H_{SE} is defined at the boundary $\mathscr{B}$. Given separability between S and E, we can, without loss of generality, write H_{SE} as:

$$H_{SE} = \beta_k k_B T_k \sum_{i}^{N} \alpha_i^k M_i^k, \tag{3}$$

where the index $k = S$ or E, k_B denotes Boltzmann's constant, T is temperature, the M_i^k are N mutually-orthogonal Hermitian operators with eigenvalues in $\{-1, 1\}$, the $\alpha_i^k \in [0, 1]$ are such that $\sum_i^N \alpha_i^k = 1$, and $\beta_k \geq \ln 2$ is an inverse measure of k's thermodynamic efficiency that depends on the internal dynamics H_k; see Fields and Marcianò (2020); Fields, Glazebrook and Marcianò (2021); Addazi et al. (2021); Fields, Glazebrook and Marcianò (2022a); Fields et al. (2022) for further motivation and details of this construction and Fields, Glazebrook and Marcianò (2022b) for a pedagogical review. Explicit time (t) dependence can be introduced by making the α_i^k time-dependent, i.e. functions $\alpha_i^k(t)$. This interaction $H_{SE}(t)$ can, again without loss of generality, be visualized as shown in Fig. 1; a fully-general control-theoretic model of $H_{SE}(t)$ is provided in Fields et al. (2023). For each of $k = S$ or E, the operators M_i^k each act on one of N qubits to either measure, or dually (Pegg Barnett and Jeffers, 2002) prepare, its state. These qubits can be regarded as constituting a Hilbert space $\otimes_{i=1}^N q_i$ that characterizes $\mathscr{B}$; note that this Hilbert space is ancillary to U, i.e. $\mathscr{B}$ is not a physical component of either S or E. The results of all N measurements by system k is an N-bit encoding of the current measured eigenvalue, for k, of H_{SE}, i.e. the current value of energy transferred to k by the interaction; the results of all N preparations by system k is an N-bit encoding of the current prepared eigenvalue, for k, of H_{SE}, i.e. the current value of energy transferred from k by the interaction. These statements imply an elapsed, system-relative time interval dt_k between measurement and preparation; where the time t_k is the fundamental time unit for system k (Fields, Glazebrook and Marcianò, 2021; Fields et al., 2022; Fields, Glazebrook and Marcianò, 2022b).
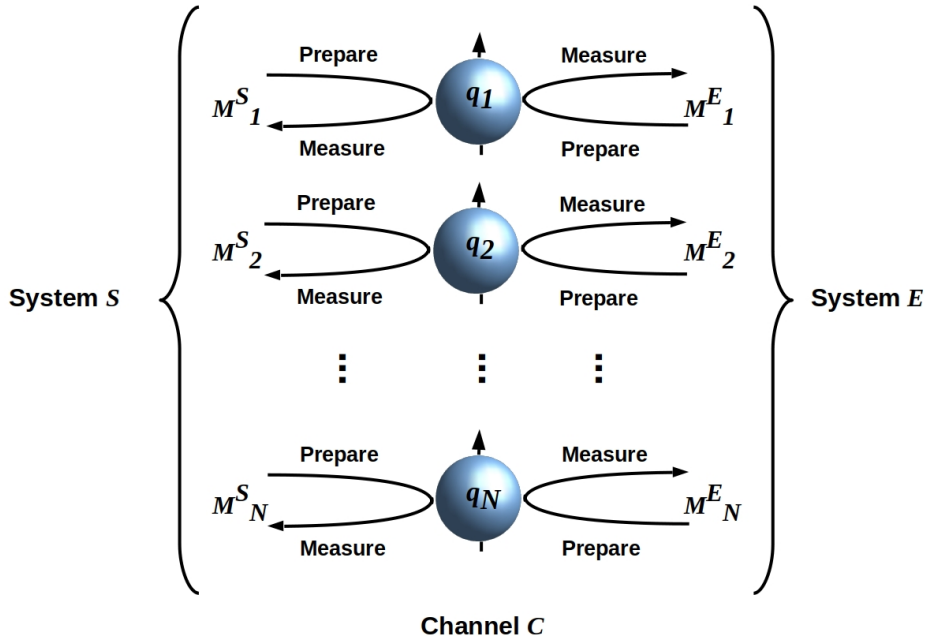
Figure 1: The action of H_{SE} given by Eq. (3) can be realized by the alternating measurement and preparation actions of each of the M_i^k on a qubit q_i. Qubits are depicted as Bloch spheres (Nielsen and Chuang, 2000). There is no requirement that S and E share preparation and measurement bases, i.e. local definitions of the z axis for each qubit. The qubits together form a communication channel C that is implemented by the boundary $\mathscr{B}$ separating S from E. Adapted from Fields and Marcianò (2020) Fig. 1, CC-BY license.

Under the conditions above, the boundary $\mathscr{B}$ is a holographic screen separating S from E: it encodes precisely the information that S can obtain about E and vice-versa. As such, it performs the functions of an MB. Hence any two physical systems which have a separable joint state, i.e. an interaction that can be represented by Eq. (3), are separated by an MB. Note that this characterization of $\mathscr{B}$, like the definition of an MB given in §2.1 above, makes no reference to 3d space; indeed it is completely topological. Note moreover that both the boundary $\mathscr{B}$ and the interaction Eq. (3) are completely invariant under further decompositions of either S or E. This follows solely from the additivity of H_U, or alternatively, the associativity of the tensor product. The internal interaction H_E, for example, can be re-written as $H_E = \sum_i H_{E_i} + \sum_{ij} H_{E_i E_j}$ without altering the equation $H_U = H_S + H_E + H_{SE}$ or the definition of $\mathscr{B}$ in any way. We will see in §3.3 below that S's identification of, and interaction with, specific "objects" or "systems" embedded in E is completely independent of any assumptions about the decomposition of E. We will then

10

consider in §4.2 how using multiple boundaries that enclose multiple hierarchical "levels" of a system of interest allows us to track system growth and development.

We are now in a position to state the FEP for generic quantum systems. As all that S can observe about E are the states of the q_i on $\mathscr{B}$ as E has prepared them, S's uncertainty about E (i.e. S's VFE) is solely uncertainty about the (next) states of the q_i. This uncertainty is reduced to zero if S's measurements of the q_i yield the same values that it, S, prepared; this will occur as $M_i^S \to M_i^E$ for all i. If $M_i^S = M_i^E$ for some i, the entanglement entropy $\mathcal{S}(SE)$ becomes non-zero; in the limit at which $M_i^S = M_i^E$ for all i, S and E become fully entangled (Fields, Glazebrook and Marcianò, 2021; Fields et al., 2022; Fields, Glazebrook and Marcianò, 2023). The limit at which VFE is minimized to zero is, therefore, the limit in which separability between S and E is fully lost. When formulated for generic systems, therefore, the FEP is the classical limit of the Principle of Unitarity, which requires the time evolution of closed systems to preserve information, and hence to asymptotically approach pure, fully-entangled states. Note that this latter principle applies to U, which by comprising "everything" of interest is, by definition, a closed system as noted earlier; its components S and E, which interact with each other, are open. The FEP applies, therefore, to all systems to which quantum theory applies, i.e. to all physical systems, provided only that they are separable from their environments. In this case, what is of interest is not *that* some system behaves in accord with the FEP, but rather *how* it does so. As Ramstead et al. (2022) emphasize, what can be tested experimentally are the proposed answers to such *how* questions. The remainder of this paper discusses in detail how a system can behave in accord with the FEP, in the specific case of living systems undergoing regulative development.

3 Informational symmetry is consistent with cognitive asymmetry

3.1 Physical interactions are informationally symmetric

As discussed in connection with Eq. (1), and as clear from the S-E exchange symmetry of Eq. (3) or Fig. 1, the FEP applies equally to any system and its environment (see also the discussion of this point in Fields et al. (2022)). The interaction H_{SE} as defined by Eq. (3), and its implementation on the inter-system boundary $\mathscr{B}$ are, in particular, informationally symmetric. Each "cycle" of interaction between S and E consists, in the representation of Fig. 1, in the encoding by E of an N-bit "message" on the qubits q_i, followed by the measurement, by S, of the q_i to produce an N-bit outcome, followed by preparation of the q_i by S and measurement of the q_i by E (Fields and Marcianò, 2020). This process effectively transfers N classical bits from E to S and then back to E. Whether the *values* of these bits are faithfully transferred between S and E – whether S will measure a value '+1' if E prepares a value '+1' – depends on whether S and E employ the same local reference frame – in this case, the same local z axis – for preparation and measurement. In general, they will

not; "free choice" of reference frame is a requirement of separability (Fields, Glazebrook and Marcianò (2022b); see also Conway and Kochen (2009) for a general discussion of free choice in this context). As we will see in §3.3 below, the local reference frames employed by systems to prepare and measure the qubits comprising their boundaries largely determine how they are able to interact with their environments, and hence how their behavior complies with the FEP. Before discussing this, however, it is necessary to understand the most fundamental aspect of system-environment interaction, thermodynamic exchange.

Information and thermodynamic energy are interconvertible by Landauer's principle (Landauer, 1961; 1999; Bennett, 1982), which sets a lower bound of $\ln 2 k_B T$ per bit on thermodynamically irreversible processes, e.g. bit erasure. Direct information-to-energy conversion in accord with this limit has been demonstrated experimentally (Toyabe et al. 2010; Bérut et al, 2012; see Parrondo, Horowitz and Sagawa, 2015 for review). Any physical interaction that irreversibly transfers information, therefore, must be accompanied by a transfer of thermodynamic energy. Processing information to obtain an answer that has an effect on subsequent behavior, i.e. irreversible classical computation (Horsman et al., 2014), has a finite free energy cost, again given by Landauer's principle. Hence any system that uses information from its environment to alter its behavior – including, clearly, any living system – must devote some of the information/energy obtained from its environment to the free-energy cost of classical computation. The amount of energy obtained from the environment thus provides a strict upper bound on the amount of classical computation (in bits per unit time) a system can perform (Fields and Levin, 2021). We can, therefore, represent any system-environment interaction as a bidirectional exchange of information across a boundary $\mathscr{B}$ in which some of the information exchanged is devoted, by each party, to the free-energy requirements of classical computation as illustrated in Fig. 2. Photosynthetic systems provide a straightforward example: some fraction of the photons received by such systems is allocated to electron transfer processes that provide free energy (e.g. ATP) to molecular pathways that process information from the environment and drive actions on the environment. The bits allocated to supplying the free-energy requirements of classical computation – e.g. those specifying photons of useable frequency in photosynthetic systems – are "uninformative" in that their specific values (i.e. +1 or -1) are irrelevant to their per-bit energetic value (at least $\ln 2 k_B T$); hence they collectively constitute an "uninformative sector" of $\mathscr{B}$, the extent of which depends on the thermodynamic efficiency of the system employing them (see Fields et al. (2022) for further discussion). The remaining bits, if any, collectively constitute the "informative sector" of $\mathscr{B}$; only these bits can provide inputs to, or encode outputs from, nontrivial computational processes. As the definitions of these sectors depend on efficiency (β_k in Eq. (3)), they will in general differ between S and E.

12

a)

$\mathscr{B}$

S E

Interaction H_{SE} defined here.

b)

Sensation (state s)

Free energy source, waste heat sink (uninformative bits)
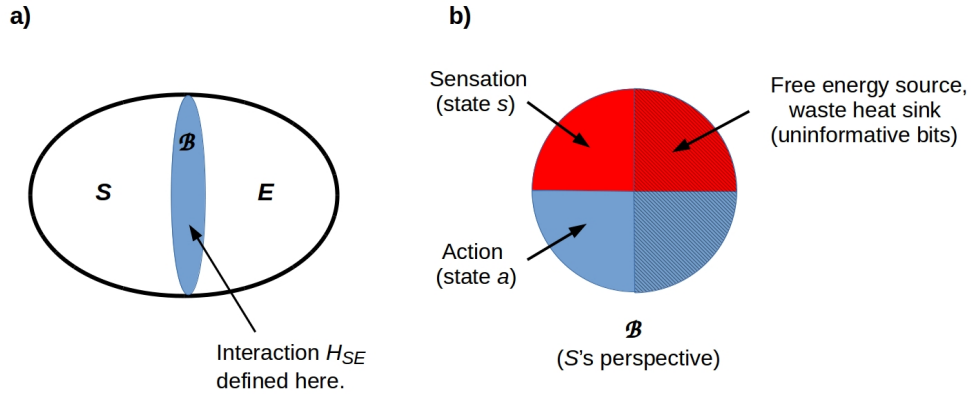
Action (state a)

$\mathscr{B}$ (S's perspective)

Figure 2: a) A system S interacts with its environment E; the Hamiltonian operator H_{SE} representing the interaction is defined at the boundary $\mathscr{B}$ that functions as an MB statistically separating S from E. The environment E comprises everything that is not the system S. b) When the boundary $\mathscr{B}$ is viewed from S's perspective, in can be divided into incoming (red) and outgoing (blue) bits. Some of these bits (gray shading) must be allocated to the thermodynamic functions of free-energy input and waste-heat dissipation; these bits are uninformative, i.e. unavailable for computational processing. The remaining bits are available as computational inputs (sensations) and outputs (actions). The situation is the same from E's perspective, though the allocation of bits to informative and uninformative sectors may be different. Adapted from Kuchling, Fields and Levin (2022), CC-BY license.

The informational symmetry of physical interactions has an important, but often neglected, consequence in the context of the FEP: the environment E of any system S must itself be considered an agent. As discussed in connection with Eq. (1), E's VFE quantifies its uncertainty about S's behavior. The FEP requires that E engages in active inference, i.e. that it acts on its environment – the system of interest S – so as to minimize the VFE that it, E, measures at its MB. In the formulation employed here, the MB of each system is just the (state space) boundary $\mathscr{B}$ between them; hence the FEP requires that any pair of interacting systems behave in such a way that they both minimize the VFE that they measured at their mutual boundary. In the limit in which the S-E interaction is purely thermodynamic, i.e. neither S nor E engages in any classical computation, both minimize measured VFE by approaching thermodynamic equilibrium. If either S or E engages in classical computation, VFE-minimizing solutions are in general not at thermodynamic equilibrium (Fields, Glazebrook and Marcianò, 2021); to the extent that S and E can be approximated by classical random dynamical systems, these solutions are NESS solutions for both S and E (see Friston (2019); Ramstead et al. (2022) for extensive details). Such solutions respect the 2nd Law from the perspectives of both component systems: S and

E each absorb free energy from and exhaust waste heat into their interaction partner, i.e. E and S respectively. Hence each system "sees" the entropy of its interaction partner increase. The (classical) entropy of the total system SE is its total information content, which is strictly conserved and hence can be rescaled to zero (cf. Tegmark, 2012 for a discussion of entropy in this bipartite-decomposition setting). This rescaling of the total entropy has no effect on the local entropies measured by either S or E.

3.2   Informational symmetry is consistent with thermodynamic asymmetry

Let us now assume that S is capable of nontrivial information processing. In this case, S implements some nontrivial function $f : s \to a$, where s is the instantaneous state of its informative (unshaded in Fig. 2) input sector and a is the instantaneous state of its informative output sector. If S is considered as a random dynamical system, the FEP requires this function to be, on average, either a gradient descent or an (internal) solenoidal flow on the statistical manifold that defines S's measured VFE as a function of E's state (Friston, 2019; Ramstead et al., 2022). In the current formulation, the function f is a (in general quantum) computation implemented by the internal interaction H_S and must, on average, be VFE non-increasing for S.

The internal dynamics H_S must respect the thermodynamic asymmetry, for S, between the informative (s, a) sector of $\mathscr{B}$ and the uninformative, thermodynamic sector. The internal dynamics H_E of E, however, are by construction conditionally independent of H_S. Hence if E also implements nontrivial information processing, there is no requirement, and indeed no expectation, that the boundary between informative and uninformative sectors for E aligns with that assigned by H_S. Nor is there any requirement or expectation that S and E have the same, or even similar, efficiencies of free energy usage or waste heat dissipation. What one interaction partner treats as an informative signal, either incoming or outgoing, can be treated by the other as uninformative "noise," i.e. as waste heat output or free energy input. Meaningful communication across $\mathscr{B}$ requires, first and foremost, overcoming this thermodynamic asymmetry. As we will see, morphology – effectively, the addition of geometric degrees of freedom to $\mathscr{B}$ – provides one way of doing this.

3.3   State measurements are reference-frame dependent

To understand what S can measure, and hence what S can "know" about E, it is useful to reflect on how we make measurements, as scientists in laboratories. Consider measuring lengths with a meter stick that has 1 mm resolution. The meter stick physically encodes a length standard – the meter – and also physically encodes, by means of permanent, unmovable tick marks, the possible outcome lengths that it can measure, each separated from the others by half of the measurement resolution. The length standard – the meter – encoded by the meter stick is arbitrary, but by encoding it, the meter stick allows measurements of

different objects at different times to be compared. The comparison process depends on two assumptions: 1) that the length of the meter stick, and hence what "1 meter" means, does not change, and 2) that the number and positions of the tick marks, traditionally called "pointer positions," do not change. Physically encoding these two assumptions makes the meter stick a *reference frame* (RF); it is useful as a reference to which other lengths may be meaningfully compared. An RF is simultaneously a physical object and a semantic object: it assigns to a measurement not just an outcome value, but also a meaning.

Human observers make ubiquitous use of reference frames, including meter sticks, clocks, and all forms of laboratory equipment, but also such non-artifacts as the diurnal cycle and the Earth's gravitational and magnetic fields. Any object that is associated, via some process of calibration, with a standard can serve as an RF. Making use of an external RF, however, always requires observation, and these observations must be mutually comparable to be physically meaningful. Hence employing an external RF for a degree of freedom such as length or time requires encoding an internal RF, including an internal standard, for that degree of freedom. Internal "standards" are fixed default values, expectation values, or set points – in dynamical systems language, NESS solutions or attractors. All internal processes that have such set points can be considered internal RFs. Humans and other organisms encode internal spatial, temporal, vibrational, chemical, and electromagnetic RFs, with RFs for the cell cycle, membrane voltage, and various chemical concentrations being some of the most ancient (see Fields and Levin (2020c); Fields, Glazebrook and Levin (2021) for explicit examples). Internal RFs solve the "problem of meaning" (Froese and Taguchi, 2019) for organisms: they render an organism's measurements mutually comparable, and hence consistently actionable. To the extent that they contribute to homeo/allostasis, they are essential for survival, as indeed recognized by the classical formulation of the FEP (Friston, 2019).

In the language employed here, any internal RFs of a system S are implemented by H_S. As H_S cannot be determined given only the interaction H_{SE} with some E, internal RFs cannot, in principle, be fully characterizable by external observations. This is, in fact, well-known independently of the current considerations: all RFs are physical systems that must, at least at microscopic scales, be considered quantum systems. They are, therefore, quantum RFs (QRFs; Aharonov and Kaufherr (1984), Bartlett, Rudolph and Spekkens (2007)) that encode "nonfungible" information, that is, information that cannot be fully specified by any finite string of classical bits (Bartlett, Rudolph and Spekkens, 2007). It is somewhat ironic that this notion of non-fungibility was discovered by quantum theorists, as classical physics assumes that physical systems are characterized by real numbers, which are not in general representable as finite bit strings. Nonfungibility has the consequence that QRFs cannot be shared by sharing bit strings, i.e. by classical communication, but only by being transferred as unique physical systems (Bartlett, Rudolph and Spekkens, 2007; Fields and Marcianò, 2019). Internal QRFs are, therefore, not sharable in principle. Indeed in the limit in which two distinct systems approach implementing identical internal QRFs, they become entangled and lose their identities as distinct systems; this is precisely the limit $M_i^S = M_i^E$ for all i at which the time evolution under the FEP becomes unitary (Fields

et al., 2022; Fields, Glazebrook and Marcianò, 2023). Hence while any QRF can be given a classical description based on finite-resolution observations, e.g. by a DNA sequence, a pathway diagram, or a parameterized differential equation, such maps can never completely characterize the (nonfungible) territory.

The information encoded by living systems is, in practice, nonfungible in a very clear sense – it typically cannot be obtained, even at low resolution, without destructive measurements that breach the system's MB and thus interfere with the internal dynamics, and hence with the very internal information that is of interest. At the cellular level, set points that define reference values are encoded in memory structures spanning the scales from that of DNA sequences, to intracellular concentration gradients, to cytoskeletal and membrane organization. Measures such as DNA sequencing, RNA profiling, and most biochemistry are irreversibly disruptive and obtain only snapshots of the cellular state. The processes that implement cellular QRFs typically involve multiple components, are often tightly coupled to other processes, e.g. by common second messengers such as Ca^{2+}, and may require quantum coherence as a resource to achieve thermodynamic feasibility (Fields and Levin, 2021); hence no measurements can be guaranteed to be free of side effects. The "measurement effect" that besets not only cell biology, but all of the life sciences, is thus not merely an analog of a quantum principle, but rather direct evidence of the non-fungible nature of biological systems.

The language of QRFs – nonfungible processes coupled to nonfungible standards – allows us to distinguish three types of measurements that a system S could, in principle, make. The first is to implement a QRF $\mathbf{Q}$ that acts on the entire input component $\mathscr{B}_{inf}(\text{in})$ of the informative sector $\mathscr{B}_{inf}$ of its boundary, as shown in Fig. 3a. The action of $\mathbf{Q}$ yields an encoding q of the instantaneous state of $\mathscr{B}_{inf}(\text{in})$. As the number of bits encoded on $\mathscr{B}$, or on any sector of $\mathscr{B}$, is proportional to its area, the free energy cost of implementing $\mathbf{Q}$ on $\mathscr{B}_{inf}$ can be written as:

$$E_{\mathbf{Q}} = A_{inf}(\text{in})\beta_{\mathbf{Q}}k_BT \tag{4}$$

where $A_{inf}(\text{in})$ is the area of $\mathscr{B}_{inf}(\text{in})$, $\beta_{\mathbf{Q}} \geq \ln2$ is a factor measuring the energetic efficiency of $\mathbf{Q}$, k_B is Boltzmann's constant and T is temperature (Fields, Glazebrook and Marcianò, 2021).
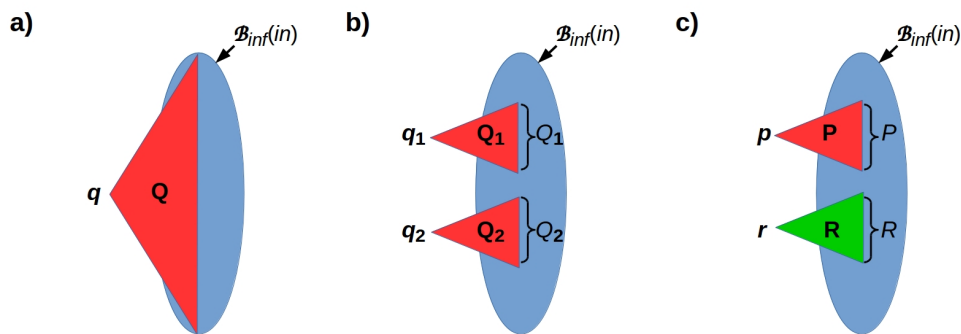
Figure 3: a) A system S implements a QRF **Q** that acts on the entire input component $\mathscr{B}_{inf}(\text{in})$ of the informative sector $\mathscr{B}_{inf}$ of its boundary, yielding an encoding q of the state of $\mathscr{B}_{inf}(\text{in})$. b) S implements QRFs **Q**$_1$ and **Q**$_2$ that act only on sectors Q_1 and Q_2 of $\mathscr{B}_{inf}(\text{in})$, yielding encodings q_1 and q_2 of the states of those sectors. c) S implements QRFs **P** ("pointer") and **R** ("reference") that act on sectors P and R of $\mathscr{B}_{inf}(\text{in})$, yielding encodings p and r of the states of those sectors, subject to the restriction that r remains fixed. Diagrams of similar form can be drawn using classical descriptions of the QRFs; standard intracellular signalling pathway diagrams, for example, have the form of Fig. 3b or c.

Assuming fixed efficiency $\beta = \beta_{\mathbf{Q}}$, S can save free energy by measuring only particular sectors of $\mathscr{B}_{inf}(\text{in})$, as shown in Fig. 3b. The state information obtained is, in this case, limited to the sector states q_1 and q_2; this limitation reflects the inevitable trade-off between energy expenditure and information gain. From a computational perspective, however, S can now measure, using further specialized QRFs, the correlation $< q_1, q_2 >$ and the conditional dependencies $(q_1|q_2)$ and $(q_2|q_1)$. These further computations enable basic logic functions (AND, OR, NOT) and, given a memory for prior probabilities, Bayesian inference (Fields and Glazebrook, 2022). Hence sector measurements enable context-dependent, explicitly hierarchical computation. A general model of the control-flow process required to deploy distinct QRFs in a context-dependent way is provided in Fields et al. (2023).

A specialized form of sector measurement is shown in Fig. 3c. Here the sector R is required to maintain a fixed state r; hence measured states of the sector P have the form $p = (p|r)$. The disjoint union $P \sqcup R$, in this case, functions as an external RF with fixed reference component R and pointer component P; the "state of interest" of PR is its "pointer" state p. All perceivable objects, or in the language of Friston (2019), all time-persistent things, detectable by S have this composite form: the invariant reference sector R allows identification of the object as "the same thing" that was measured earlier, while the variable pointer sector P displays the state of interest. Again items of laboratory apparatus provide a canonical example since they must be distinguished from their surroundings by measuring

17

such observables as size, shape, color, and position before the informative "pointer" readouts can be determined.

3.4   Hierarchical measurements are ubiquitous at all scales

Turning now to biology, it is clear that measurements of the forms shown in Fig. 3b and 3c are ubiquitous, while measurements of the form shown in Fig. 3a can occur only in the limit of no post-measurement classical processing. Measuring the entire detectable state of the environment at every instant leaves no space on $\mathscr{B}$ to support thermodynamic exchange; organisms instead focus their resources on what is salient and significant, deploying different sector-specific QRFs at different times. Allocating resources to the salient and significant is the function of attention systems (see e.g. Burgoyne and Engle (2020) for a recent review and Fields, Glazebrook and Levin (2021) for discussion in the current context).

While both object perception and comparative measurements are familiar in animals with large brains, they are less well characterized in smaller-brained or aneural animals, plants, and unicellulars. All organisms capable of identifying and behaving in state-specific ways toward mates or conspecifics, however, are capable of some level of object identification, and hence of Fig. 3c measurements. This capability extends, therefore, into the microbial world. All organisms capable of conditioning their behavior on the values of two variables simultaneously, e.g. on sugar and salt concentration, or on ambient light and the availability of water, are capable of Fig. 3b measurements. This capability extends, therefore, even deeper into the microbial world, and to our knowledge characterizes all of life. The same considerations apply at intracellular scales. Enzymes that require cofactors or phosphorylation as well as a substrate are performing Fig. 3b measurements on their environments, as are membrane-bound channels, pumps, and receptors that are sensitive to both local membrane voltage and binding of GTP or other regulatory molecules. RNA polymerases that bind DNA and then scan for specific transcription-initiation sites (Kuehner and Brow, 2006) are performing Fig. 3c measurements on their environments: first identifying a DNA molecule as an object, and then detecting a particular state – a particular base-sequence motif – of that object. Spatially-organized intracellular pathways are increasingly recognized as performing complex, multi-input computations, both in neurons (Gidon et al., 2020) and in non-neural cells (Kramer, del Castillo and Pelkmans, 2022), as long argued by the basal cognition movement on the basis of both biochemical and behavioral data (Maturana and Varela, 1980; Pattee, 1982; Stewart, 1996; di Primio, Müller and Lengeler, 2000; Lyon, 2015; Baluška and Levin, 2016; Baluška and Reber, 2019; Levin, 2019; Lyon, 2020).

3.5   Hierarchies of QRFs are MCAs

We have, in the above, focused on measurements, i.e. on processing the bits encoded on the incoming side of the informative sector $\mathscr{B}_{inf}(\text{in})$ of the S-E boundary. Actions on

E can, however, be viewed as "preparations" of the bits encoded on the outgoing side of the informative sector $\mathscr{B}_{inf}$(out). Preparing or setting a bit value is dual, as a process, to measuring its value (Pegg, Barnett and Jeffers, 2002). These processes are, therefore, informationally symmetric; indeed any QRF can be viewed as preparing, as well as measuring, the values of some degree(s) of freedom. A meter stick, for example, can be used to mark 30 cm lengths as well as to measure them. This informational symmetry is manifest when QRFs are represented as category-theoretic structures – formally, limits and colimits – of networks of bidirectional operators on single strings of bits, a representation that is provably completely general (Fields, Glazebrook and Marcianò, 2022a).

If all QRFs are informationally symmetric, the mapping $f : s \to a$ implemented by the internal dynamics H_S must be informationally symmetric. The informational symmetry of H_{SE} then requires that the thermodynamic functions of free energy acquisition and waste heat dissipation must be informationally symmetric. The dynamics H_S can, therefore, be divided into two informationally symmetric flows, one of which transfers free energy to and dissipates waste heat from the other. We can, therefore, redraw Fig. 2b as Fig. 4. The analogy with information and free energy flows in chemical reactions is obvious.
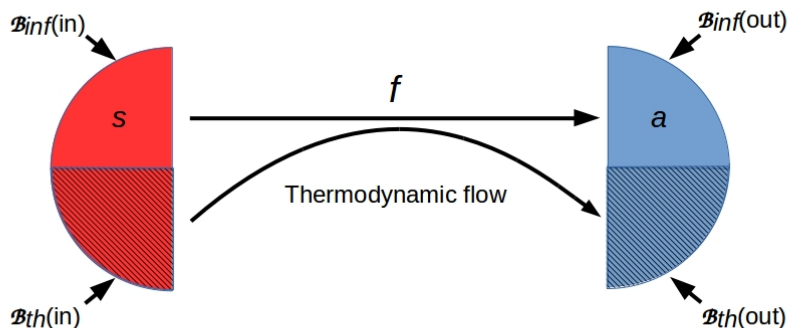

Figure 4: The mapping $f : s \to a$ implemented by S's QRFs is supported by an informationally-symmetric thermodynamic flow from the uninformative (thermodynamic) sector $\mathscr{B}_{th}$(in) to the corresponding output sector $\mathscr{B}_{th}$(out).

The mapping $f : s \to a$ is, at each instant, implemented by some hierarchy of QRFs on both input and output sides; we have previously showed, as an example, how cortical neurons implement QRF hierarchies (Fields, Glazebrook and Levin, 2022). Which QRFs are deployed at any instant determines a measurement context (Fields and Glazebrook, 2022; Fields et al., 2022); switching between contexts is a metaprocessing function which, as a component of $f : s \to a$, must itself be implemented by a (fixed) QRF hierarchy (Kuchling, Fields and Levin, 2022). Each QRF in the hierarchy is a self-contained computational

system with its own inputs, outputs, power supply, and semantics. The QRF hierarchy is, therefore, a canonical MCA.

4 Biological self-organization is thermodynamically driven

4.1 The "environment" is an active agent

It is well known that simple physical systems can self-organize complex structures when subjected to mechanical and thermodynamic forcing; thunderstorms provide the most familiar example. Turing (1953) introduced such processes to the study of morphogenesis; see Kondo and Miura (2010); Morelli et al. (2012) for reviews. In practice, such mechanisms are generally conceptualized as occurring inside the system of interest, e.g. inside a cell, or inside a developing organism. By considering the "environment" to be outside of the system undergoing pattern formation, the role of the environment as a thermodynamic agent is minimized. This renders such pattern formation models *prima facie* consistent with the neo-Darwinian, genome-focused approach to morphogenesis (e.g. Monod, 1972; Dawkins, 1984; Michod, 1999) with its idea that "[d]evelopmental biology can be seen as the study of how information in the genome is translated into adult structure, and evolutionary biology of how the information came to be there in the first place" (Szathmáry and Maynard Smith, 1995, p. 231).

By focussing attention on processes occurring at system-environment boundaries, the FEP framework allows us to examine the role of the "environment" of a process not at some pre-determined scale, but at the scale of the actual dynamics of interest. As emphasized by Friston (2019), *all* physical systems that persist through time can be understood as self-organizing from the perspective of the FEP. We are now in a position to use the conceptual tools provided by the FEP to understand self-organization both within a given scale and, more interestingly, as a thermodynamically-driven process that generates complexity at progressively larger scales. This will allow us to connect origin-of-life models to evolutionary and developmental models within a single formal framework, and to understand how the transition from protocellular to cellular and then multicellular systems is driven by the FEP.

The FEP has already been shown to provide a generic model of within-scale self-organization for systems with fixed MBs (Kirchhoff et al., 2018; Friston, 2019). To briefly summarize, VFE minimization corresponds to maintaining a state close to the NESS. This is achieved when internal processes – i.e. the internal dynamics H_S – predict the environment's actions on the MB sufficiently accurately to maintain the integrity of the MB and hence the conditional independence of internal states. This is referred to as "self-evidencing" (Friston, 2019) and corresponds, in biological systems, to the maintenance of homeo/allostasis (Friston, 2013). To place this in the current language, the internal dynamics H_S, and in particular, the QRF hierarchy that it implements, can in this case be regarded as a GM of the behavior of E, specifically, of the action of E on the MB.

We have also shown previously how producing "copies" of itself that cluster in the local environment is a viable strategy for a system to reduce VFE (Fields and Levin, 2019). The copies shield the system from the open environment, reducing its unpredictability. The behavior of the copies is similar to the behavior of the system, thus increasing predictability. While this model was formulated for biological cells and shows how multicellularity can be advantageous from the perspective of the FEP, it applies at larger scales as well, with ethnic, linguistic, or religious communities and social-media "echo chambers" as obvious examples.

Here we take a different, but complementary perspective. The "environment" surrounding an open system is typically treated as passive: as a thermodynamic or material-exchange resource, an ambient field, or simply a heat bath. In the FEP literature, the environment is often just a source of uncertainty. For biological systems, this is clearly unrealistic: the environment of any organism is itself an evolved structure that includes products manufactured by other organisms, e.g. oxygen, soil, and more recently, human artifacts. When all of life is considered one developing system, "the environment" becomes that system's stigmergic memory (Fields and Levin, 2020b). At the scale of a single organism, the immediate environment consists largely of other organisms, both conspecifics and others. Similarly, the environment of a biomolecule consists largely of other biomolecules, the environment of a population consists largely of other populations, etc. These "others" are active agents pursuing their own agendas, as made explicit in game-theoretic models.

The FEP not only allows, but when viewed in full generality, requires this game-theoretic perspective. The FEP applies to all systems with MBs, and describes all such systems as VFE-minimizing agents. The environment of any system shares an MB with that system, as made explicit in Fig. 2a. The environment E of any system S is, therefore, a VFE-minimizing agent. The sole source of VFE for E is S; any generative model implemented by E is, therefore, a generative model of S's actions on its MB. To assume that E's generative model is random – that E functions only as a heat bath or noise source – is thus to assume a very special case, one that is largely irrelevant to biology. The FEP thus extends and generalizes an insight of Lovelock and Margulis (1974), Maturana and Varela (1980), Rosen (1986), and many others: realistic environments are active agents, just as more typical systems of interest are. The capabilities of the environment as an agent depend only on its dynamics H_E and hence are completely independent of decomposition as noted in §2.2; in particular, they are independent of its description by the system of interest using its QRFs, i.e. in terms of the objects perceived by the system.

4.2 The environment acts so as to increase the system's predictability

When we consider E to be an active agent, the goal of its actions becomes clear: E acts to decrease its measured VFE; hence it acts to increase the predictability, by its generative model, of S's behavior. In particular, E acts to increase the predictability of S's actions on its MB. Note that from E's perspective, S's behavior becoming increasingly predictable

corresponds to S transferring less novel information to E and hence "losing freedom." The rate of increase of S's entropy, relative to E, thus decreases, though it remains positive since S still absorbs all waste heat generated by E, as discussed in §3.1 above. The considerations of the previous section apply equally to S and to E; to understand E's actions on S, we must consider E's information-processing capabilities, i.e. its hierarchy of QRFs. It is, in particular, important to understand whether E processes S's actions on $\mathscr{B}$ as informative inputs and vice versa (Fields et al., 2022).

A model in which S inserts "copies" of itself into E (Fields and Levin, 2019) is effectively a model in which E deploys QRFs that "make sense" of S's actions, as illustrated in Fig. 5a – 5c. From S's perspective, the copies are components of E – or more precisely, the behaviors of the copies are components of the behavior of E, as measured by S at its MB – that are at least partially predictable and hence "make sense" to S. As the number of copies increases from zero (Fig. 5a) to one (Fig. 5b) to many (Fig. 5c), the behavior of E as a whole becomes progressively more predictable by S. Hence it is advantageous, from an FEP perspective, for S to insert copies of itself into E (Fields and Levin, 2019). The same, however, is true for E: as E gains copies of S and hence incorporates their QRFs, the behavior of S becomes more predictable for E. As noted earlier, this symmetric dynamic becomes obvious when we consider the QRFs that implement human language understanding or other sociocultural practices. A company E, for example, can predict the behavior of its customer S much better if it includes employees C that speak S's language.
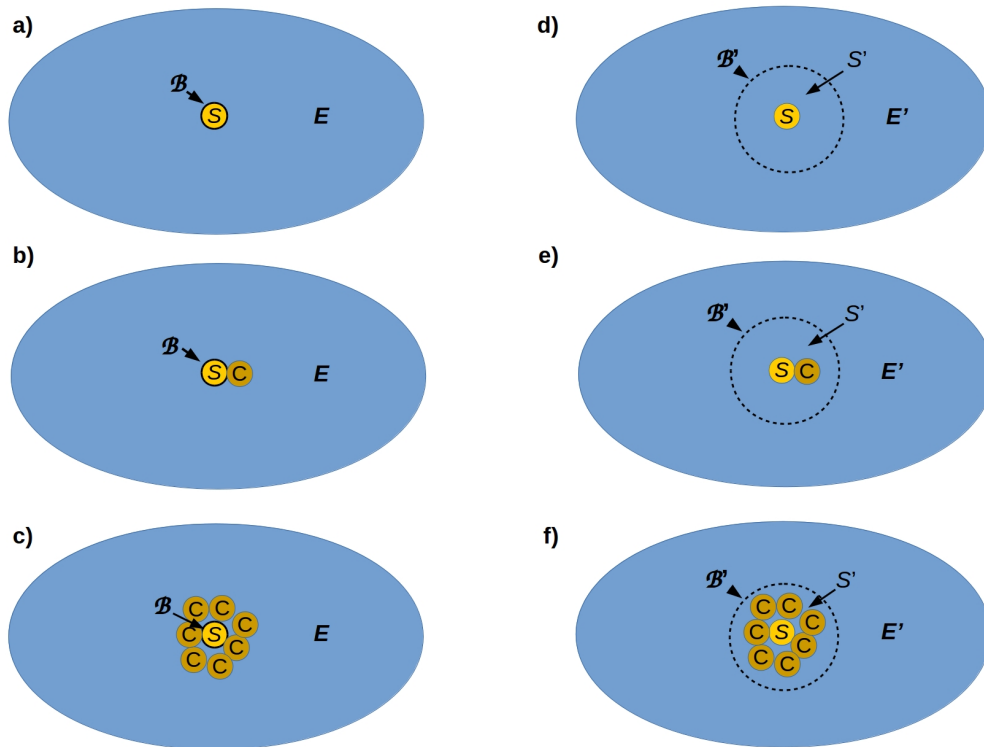
Figure 5: a) A system S interacts with its environment E as in Fig. 2a. b) S inserts a "copy" C of itself into E. The copy brings new QRFs to E, increasing E's predictive power. c) S inserts multiple copies of itself into E, as in the model of Fields and Levin (2019). d) A boundary $\mathscr{B}'$ can be drawn anywhere in E. Provided the states inside and outside $\mathscr{B}'$ are mutually conditionally independent (i.e. $\mathscr{B}'$ functions as an MB), drawing this new boundary defines new systems S' (inside $\mathscr{B}'$ and containing S) and E' (outside $\mathscr{B}'$, the remainder of E). e) S adds a copy C of itself to S'. f) S adds multiple copies of itself to S'.

We have thus far considered the interaction H_{SE} and the boundary $\mathscr{B}$ between S and E. We are, however, free to pick any boundary in the joint system SE that we like, provided only that it functions as an MB in rendering the states of the two systems that it separates mutually conditionally independent. Let us suppose, therefore, that the boundary $\mathscr{B}'$ shown in Fig. 5d meets this condition. Drawing this boundary defines a system S' – which includes S – that is inside $\mathscr{B}'$, and an environment E' that is outside $\mathscr{B}'$. The two interact, at $\mathscr{B}'$, via an interaction $H_{S'E'}$. All of the previous considerations apply to this new interaction. We can now consider the consequences of S inserting copies of itself into S', i.e. into the interior defined by the boundary $\mathscr{B}'$, as shown in Fig. 5e and 5f. Doing this has the consequences for S's predictive capability discussed above. Here, however, we will be interested in its consequences for the predictive capabilities of S' and E'.

When the boundary is moved from $\mathscr{B}$ to $\mathscr{B}'$, some degrees of freedom of E become degrees of

freedom of S'; hence E' has fewer degree of freedom than E, and therefore less computational power than E. Any QRFs of E that directly measured bits encoded on $\mathscr{B}$, in particular, are lost in the transition from E to E'. However, because $\mathscr{B}'$ by construction functions as an MB, the states of S' are conditionally independent of the states of E'. In this case, E factors as $E = E' \cdot (S' \setminus S)$, where '$\setminus$' denotes set, or more properly state-space subtraction. Any QRFs implemented by $S' \setminus S$ that measure states of $\mathscr{B}$, therefore, can be viewed as writing their outcome values on $\mathscr{B}'$. We can, therefore, view $\mathscr{B}'$ as a coarse-graining of $\mathscr{B}$. Indeed, this construction implements the idea of "MBs within MBs" (Kirchhoff et al., 2018).

This construction underscores an important point, one relevant to some criticisms of the applicability of the FEP to biological systems. When a cell, for example, divides, it does not lose its MB. Its environment, however, changes; hence its interaction with its environment changes. The joint system comprising the cell and its daughter can also be regarded as a "system of interest" with its own MB, one that is substantially larger that the MB of the original cell. With sufficient divisions, the MB of the developing multicellular system may wholly contain the MB of the original cell, together with many other MBs surrounding pairs, triples, or other collections of component cells. Each of these MBs defines a different system, a different environment, and a different system-environment interaction. The MB of one's liver, for example, supports a very different interaction than the MB of one's brain. Understanding a whole organism's interaction with its environment – as defined at the whole organism's MB – requires also understanding the interactions defined at all of the MBs within the organism: behavioral biology requires physiology. This is not a reductionist view: understanding the interactions between the parts and their environments is a necessary, but not a sufficient, condition for understanding the interaction of the whole with its environment. It is rather an integrative, systems-biology view: the hierarchy of MBs within the organism specifies the loci of the internal interactions that both enable the whole organism's interaction with its environment and enable that interaction to be understood.

A particular example of the above construction is familiar, and has been studied in detail. Suppose S is a system of interest, E' is an observer, e.g. a human observer, and $S' \setminus S$ is an ambient field, e.g. the ambient visible-frequency photon field. In this case, the QRFs implemented by $S' \setminus S$ that measure $\mathscr{B}$ are light-scattering interactions that encode properties of S, e.g. size or shape. The scattered light impacts E' at the boundary $\mathscr{B}'$, transferring the encoded information about S to E'. The resolution of the encoding is decreased by a factor proportional to the ratio of the areas of $\mathscr{B}$ and $\mathscr{B}'$; the transfer process thus implements a coarse-graining. Clearly, this coarse-graining information-transfer process is simply the standard classical-optics mechanism of visual perception. Such environment-mediated measurement has been studied in general under the rubric of "quantum Darwinism" (Zurek, 2003; 2009). It depends critically on boundaries that impose conditional independence at both $\mathscr{B}$ and $\mathscr{B}'$.

Given this conditional independence condition, we can write $H_E = H_{E'} + H_{S' \setminus S}$, where $H_{S' \setminus S}$ implements QRFs transferring information between $\mathscr{B}$ and $\mathscr{B}'$ and $H_{E'}$ implements

QRFs outside $\mathscr{B}'$. The condition that allows drawing $\mathscr{B}'$ is separability: $\mathscr{B}'$ must function as an MB that renders E' conditionally independent of S'. The interaction between E' and S' must, therefore, be effectively classical; any QRFs of E that cross the new interaction boundary $\mathscr{B}'$ must include a classical interface – effectively, an application-programming interface (API) – at $\mathscr{B}'$ that prevents sharing quantum coherence across $\mathscr{B}'$. Under these conditions, drawing the boundary $\mathscr{B}'$ has no effect on H_E and hence no effect on E's QRF hierarchy. The component E' therefore "knows" exactly the same things about S in Fig. 5a and Fig. 5d; Fig. 5d simply makes the boundary $\mathscr{B}'$ explicit. The same equality must, in this case, hold between what E' knows about S in Fig. 5b and 5e, and in Fig. 5c and 5f. What is going on inside $\mathscr{B}'$ does not change the E'-S' interaction in any of these pairs of situations.

What, then, is the effect of the copies C filling up $S' \setminus S$ from the point of view of E'? What is the difference, for E', between Fig. 5d and 5f, or between Fig. 5a and 5c? The copies C add their own QRFs to $S' \setminus S$. These QRFs transfer information about S to E', and also transfer information about E' to S as discussed in Fields and Levin (2019). The area, and hence the coding capacity, of $\mathscr{B}'$ does not change as these additional QRFs are inserted into $S' \setminus S$. The boundary $\mathscr{B}'$ must, therefore, encode a convolution, which for simplicity we can take to simply be an average of the outcomes written by the QRFs contributed by the multiple copies C. These outcomes are likely to be similar, since the C are all copies of S, but in general they will not be identical. Hence in general, $\mathscr{B}'$ implements a further coarse-graining of information about S in Fig. 5f compared to 5d, or between Fig. 5a and 5c. Transfer of information through an ambient medium again provides a familiar example of this coarse-graining effect. Listening to a radio, or a conversation, in a quiet room can be a low-noise, high-resolution experience. Adding multiple, inexact copies of the radio, or of the conversation, inevitably increases noise and decreases fidelity.

While coarse-graining increases noise, it also decreases uncertainty if the resolution of measurements is also decreased. This is achieved by deploying QRFs that measure relatively low-resolution "macrostates" in place of high-resolution microstates (Hoel, Albantakis and Tononi, 2013; Hoel, 2017). From the perspective of E', therefore, filling up $S' \setminus S$ with copies C of S decreases VFE, and therefore increases predictive power, provided E'''s QRFs are lower-resolution than E's QRFs. This is the case whenever E's QRFs are hierarchies that cross $\mathscr{B}'$: in this case, E'''s components of the hierarchy automatically coarse-grain E's complete QRF. Neurons are structured so as to take advantage of such coarse graining (Fields, Glazebrook and Levin, 2022). Coarse-graining and hierarchical structure further increase predictive power if the copies C are diverse, as the coarse-graining imposed by $\mathscr{B}'$ will "wash out" the details written by the C more if they are writing more diverse outcomes. Hence we have a rather surprising conclusion:

> The FEP will drive the peripheral environment E' around any system S to act on S so as to enable or facilitate the insertion of diversified "copies" of S into S's immediate environment.

The FEP, in other words, drives the environment around any "interesting" system to enable

both the replication of that system and the self-organization of the replicates into a "body" surrounding the system. Self-organization is, therefore, environmentally-driven under the FEP.

4.3 Environmental driving facilitates both the origin of life and its diversification

The surprising conclusion above becomes unsurprising when we consider it in the case of either animal or plant reproduction: germ cells are intentionally enclosed in microenvironments – seeds, eggs, a uterus – that facilitate their replication and the self-organization of an embryo. The "correct" microenvironment is generally essential to the success of the self-organization process.

The same is obviously true in bioengineering and artificial life contexts; here the "environment" in the form of the human experimenter provides both the materials and the microenvironment required for self-organization. Here as in all curiosity-driven experimentation, the goal of the experimenter-as-environment is to increase future predictive power, i.e. to decrease future VFE. Both the rate and extent of the decrease depend on the complexity of the engineered system, particularly on the complexity of its QRFs. Machine learning (ML) systems, for example, are designed to generate their own QRFs that encode their accumulated experience. Systems such as AlphaFold (Jumper et al., 2021) and AlphaCode (Li et al., 2022) exhibit surprising (and surprisingly useful) behavior; this potential for surprise generates the "explanation problem" – a problem of how to reduce experimentally-elevated VFE – in AI (Arrieta et al., 2020; Samek et al., 2021). Future hybrid and chimeric systems that incorporate biological "parts" are expected to generate even more challenging explanation problems (Levin, 2022).

Bioengineering and artificial life contexts highlight the environment's ability to "take over" the task of replication, providing the copies C that are needed and sometimes actually inserting them into the immediate microenvironment of the system. In the case of kinematic replication of xenobots, for example, the environment provides the needed supply of dissociated *X. laevis* skin cells (Kriegman et al., 2021). The environment also provides the "parts" in naturally-occurring cases of affiliative aggregation, from *Dictyostelium* sporulation or the formation of multispecies microbial mats to the replication of symbiotic systems, including all holobionts (Guerrero, Margulis and Berlanga, 2013; Gilbert, 2014a; Bordenstein and Theis, 2015). As humans are holobionts, our own reproduction is environmentally assisted; while the core microbiota is transferred from the mother at birth (Gilbert, 2014b; Coscia et al., 2021), further components are added by the environment – including the nursing mother – after birth, and indeed throughout the lifespan.

Viewed more broadly, the environment provides the parts in every case of biological or biochemical replication, in the form of molecular subunits to be assembled by an essentially kinematic process. Replication of DNA – the fundamental "replicator" in the gene-centric neo-Darwinist view of both evolution and development (Dawkins, 1984) – is an

environmentally-driven process: the environment provides the nucleic acids, the enzymes, the free energy, and the biochemically and thermodynamically stabilized compartment required for the kinematic process. Why? From an FEP perspective, the environment does these things to increase future predictability. Making more of the same kind of molecule generates a more predictable future state than making a random assortment of molecules.

The environment similarly provides both parts and a stabilized microenvironment in origin-of-life models (Cornish-Bowden and Cárdenas, 2017; Bartlett and Wong, 2020). The FEP suggests that it does this for the same reason that it does this in the case of DNA replication or organismal reproduction: to increase its future predictive power. Living systems localize, organize, and coarse-grain information. From the point of view of the environment, this compartmentalization reduces VFE.

As discussed in Friston (2019) and Fields et al. (2022), the classical limit of increasing predictive power is generalized synchrony between system and environment: each predicts the other's future actions perfectly. Achieving this limit, clearly, requires evenly-matched computational power, and hence both behavioral and computational capabilities, on each side of the system-environment boundary. "Small" organisms – e.g. microbes – compensate in part for their relatively limited computational power by limiting the sizes, and thus the informational bandwidths, of their MBs. If most environmental variation is invisible, the task of predicting what is visible becomes easier. Humans represent an opposite extreme, as we continually expand the informational bandwidths of our MBs by developing new observational technologies. This activity – science – increases our VFE and hence the predictive demands placed on our cognitive systems. Rendering the environment more predictable does not, however, render it less complex; it rather requires that our predictive models – our GMs – become more complex. A limiting scenario of generalized system-environment synchrony can involve arbitrarily large (but of course finite) complexity on both the system and the environment side. Human evolution itself is a microcosm of this process: the increasing complexity of primate, and then human, social interactions are widely acknowledged to have co-evolved as an "arms race" with the increasing complexity of primate, and then human, cognition, including such cognitive complexities as language use, extensive environmental manipulation, and culture (Adolphs, 2003; Dunbar, 2003; Dunbar and Shultz, 2007; Adolphs, 2009). Predicting the outcomes of such arms races would require detailed models of both the perception/action capabilities (i.e. the QRFs) and the computational capabilities of both the components and their shared environment. This is particularly true in cases in which the multiple components of a complex system, e.g., a population or an ecosystem, may have conflicting goals.

5   Conclusion

We have seen here how the FEP provides a generic model of MCAs that applies equally to natural and engineered systems and equally to short and long timescales. Indeed the FEP erases the distinction between natural and engineered systems. Because the FEP charac-

terizes the environment of any system of interest as an agent, the environment can always be regarded as "training" or "engineering" the system. The ubiquitous role of the environment in providing the parts required for replication, whether of DNA molecules, cells, evolved or constructed organisms, or completely artificial, abiotic systems, demonstrates this engineering aspect. When the environment is seen as an engineer, it becomes clear that "self-organization" is always environment-assisted self-organization. The product of any such process, the FEP tells us, serves to decrease the environment's measured VFE, and hence in an important sense serves the environment's goals.

This view of the environment and its interactions with living systems significantly broadens the usual concept of what is "normal" or "typical" in biology. Xenobots and chimeras become exemplars, not oddities. It also becomes clear that the environment encodes "target morphology" in the form of VFE reduction criteria all the way down. In both evolution and development, and in origin of life scenarios, the environment assembles a bunch of likely parts to see what happens. Life is a successful outcome of an experiment performed by the environment.

Why would an environment assemble parts to create a living system, and then assist in its replication? The FEP suggests a simple answer: the environment is an agent that creates novelty in order to see what information it can get in return. The environment is a typical active inference agent. It behaves like any such agent behaves, limited only by the free energy it can obtain and the computational resources it can bring to bear.

This way of thinking suggests an experimental strategy that has been pursued, but never systematically: it suggests dissociating embryos or other collections of cells, of various different kinds, mixing them together in diverse, "multicultural" populations, and seeing how they behave in various environments. Can we make xenobot-like systems, for example, that are multi-origin chimeras? Can we make *de novo* symbiotic complexes, analogous to lichens, that have parts from very different lineages? Experiments along these lines would, in effect, be a kind of "recombinant biology", analogous to standard genetic engineering, but carried out with cells, not genes. The success of recombinant genetics suggests that recombinant biology may work for cells in some "right" kinds of environments. The outcomes of such experiments could substantially increase the diversity of life beyond that supplied thus far by evolution.

Acknowledgements

Conflict of interest

The authors declare no competing, financial, or commercial interests in this research.

References

Addazi, A.; Chen, P.; Fabrocini, F.; Fields, C.; Greco, E.; Lulli, M.; Marcianò, A.; Pasechnik, R. (2021) Generalized holographic principle, gauge invariance and the emergence of gravity à la Wilczek. *Front. Astron. Space Sci. 8*, 563450.

Adolphs, R. (2003) Cognitive neuroscience of human social behavior. Nat. Rev. Neurosci. 4, 165-178.

Adolphs, R. (2009) The social brain: neural basis for social knowledge. Annu. Rev. Psychol. 60, 693-716.

Aguilera, M., Millidge, B., Tschantz, A. and Buckley, C. L. (2021) How particular is the physics of the free energy principle? Phys. Life Rev. 40, 24-50.

Aharonov, Y. and Kaufherr, T. (1984) Quantum frames of reference. Phys. Rev. D 30, 368-385.

Arrieta, A. B., N. Díaz-Rodríguez, J. Del Ser et al. (2020) Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Inform. Fusion 58, 82-115.

Baluška, F. and Levin, M. (2016) On having no head: Cognition throughout biological systems. Front. Psychol. 7: 902.

Baluška, F. and Reber, A. (2019) Sentience and consciousness in single cells: How the first minds emerged in unicellular species. BioEssays 41, 1800229.

Bargh, J. A. and Ferguson, M. J. (2000) Beyond behaviorism: On the automaticity of higher mental processes. Psychol. Bull. 126, 925-945.

Barrière, A. and Bertrand, V. (2020) Neuronal specification in C. elegans: combining lineage inheritance with intercellular signaling. J. Neurogenet. 34, 273-281. https://doi.org/10.1080/01677063.2020.1781850

Bartlett, S. D., Rudolph, T. and Spekkens, R. W. (2007) Reference frames, super-selection rules, and quantum information. Rev. Mod. Phys. 79, 555-609.

Bartlett, S. and Wong, M. L. (2020) Defining Lyfe in the Universe: From three privileged functions to four pillars. Life 10, 42.

Bennett, C. H. (1982) The thermodynamics of computation. Int. J. Theor. Phys. 121, 905-940.

Bérut, A., Arakelyan, A., Petrosyan, A., Clberto, S., Dllenschneider, R. and Lutz, E. (2012) Experimental verification of Landauer's principle linking information and thermodynamics. Nature 483, 187-189. https://doi.org/10.1038/nature10872

Biehl, M., Pollock, F. A. and Kanai, R. (2021) A technical critique of some parts of the free energy principle. Entropy 23, 293.

Birnbaum, K. D. and Alvarado, A. S. (2008) Slicing across kingdoms: Regeneration in plants and animals. Cell 132, 697-710.

Bissell, M. J.; Radisky, D. C., Rizki, A., Weaver, D. M. and Peterson, O. W. (2002) The organizing principle: Microenvironmental influences in the normal and malignant breast. Differentiation 70, 537- 546.

Bizzarri, M. and Cucina, A. (2014) Tumor and the microenvironment: A chance to reframe the paradigm of carcinogenesis? Biomed. Res. Int. 2014, 934038.

Bohr, N. (1928) The quantum postulate and the recent development of atomic theory. Nature 121, 580–590.

Bohr, N. (1958) Atomic Physics and Human Knowledge. Wiley, New York.

Bordenstein, S. R. and Theis, K. R. (2015) Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. PLoS Biol. 13(8), e1002226.

Bruineberg, J., Dolega, K., Dewhurst, J. and Baltieri, M. (2022) The emperor's new Markov blankets. Behav. Brain Sci. 45, e183.

Burgoyne, A. P. and Engle, R. W. (2020) Attention control: A cornerstone of higher-order cognition. Curr. Dir. Psych. Sci. 29, 624-630.

Clark, A. (2017) How to knit your own Markov blanket: Resisting the second law with metamorphic minds. In: Wetzinger, T., Wiese, W. (Eds.), Philosophy and Predictive Processing, Vol. 3. Frankfurt am Mainz Mind Group, p. 19pp.

Clawson, W. and Levin, M. (2022) Endless forms most beautiful 2.0: Teleonomy and the bioengineering of chimaeric and synthetic organisms, Biol. J. Linnean Soc. 2022, blac073. https://doi.org/10.1093/biolinnean/blac073

Conway, J. H. and Kochen, S. (2009) The strong free will theorem. Notices AMS 56, 226-232.

Cornish-Bowden, A. and Cárdenas, M. (2017) Life before LUCA. J. Theor. Biol. 434, 68-74.

Coscia, A., Bardanzellu, F., Caboni, E., Fanos, V. and Peroni, D. G. (2021) When a neonate is born, so is a microbiota. Life 11, 148.

Dawkins, R. (1984) Replicators and vehicles. In: Brandon, R. N. and Burian, R.M. (eds) Genes, Organisms, Populations: Controversies Over the Units of Selection. Cambridge, MA: The MIT Press, pp. 161-180.

Deutsch, D. (1997) The Fabric of Reality. New York: Penguin.

Di Paolo, E., Thompson, E. and Beer, R. (2022) Laying down a forking path: Tensions between enaction and the free energy principle. Philos. Mind Sci. 3, 2.

di Primio, F., Müller, B. F. and Lengeler, J. W. (2000) Minimal cognition in unicellular organisms. In: Meyer, J.A., Berthoz, A., Floreano, D., Roitblat, H. L. and Wilson, S. W. (Eds), From Animals to Animats. International Society for Adaptive Behavior: Honolulu, HI, USA, pp. 3-12.

Dunbar, R. I. M. (2003) The social brain: mind, language and society in evolutionary perspective. Annu. Rev. Anthropol. 32, 163-181.

Dunbar, R. I. M., and Shultz, S. (2007) Evolution in the social brain. Science 317, 1344-1347.

Durant, F., Morokuma, J., Fields, C. Williams, K., Adams, D. S. and Levin, M. (2017) Long-term, stochastic editing of regenerative anatomy via targeting endogenous bioelectric gradients. Biophys. J. 112, 2231-2243.

Farinella-Ferruzza, N. (1956) The transformation of a tail into limb after xenoplastic transplantation. Experientia 12, 304-305.

Farnsworth, D. R., Saunders, L. M. and Miller, A. C. (2020) A single-cell transcriptome atlas for zebrafish development. Devel. Biol. 459, 100-108. https://doi.org/10.1016/j.ydbio.2019.11.008

Fields, C., Fabrocini, F., Friston, K., Glazebrook, J. F., Hazan, H., Levin, M. and Marcianò, A. (2023) Control flow in active inference systems. Preprint OSF 8e4ra, DOI:10.31219/osf.io/8e4ra.

Fields, C., Friston, K., Glazebrook, J. F., Levin, M. (2022) A free energy principle for generic quantum systems. Prog. Biophys. Mol. Biol. 173, 36-59. https://doi.org/10.1016/j.pbiomolbio.2022.05.006

Fields, C. and Glazebrook, J. F. (2020) Do Process-1 simulations generate the epistemic feelings that drive Process-2 decision making? Cogn. Proc. 21, 533-553. https://doi.org/10.1007/s10339-020- 00981-9

Fields, C. and Glazebrook, J. F. (2022) Information flow in context-dependent hierarchical Bayesian inference. J. Expt. Theor. Artif. Intell. 34, 111-142 https://doi.org/10.1080/0952813X.2020.1836034

Fields, C., Glazebrook, J. F. and Levin, M. (2021) Minimal physicalism as a scale-free substrate for cognition and consciousness. Neurosci. Cons. 2021, niab013. https://doi.org/10.1093/nc/niab013

Fields, C., Glazebrook, J. F. and Levin, M. (2022) Neurons as hierarchies of quantum reference frames. BioSystems 219, 104714. https://doi.org/10.1016/j.biosystems.2022.104714

Fields, C., Glazebrook, J. F. and Marcianò, A. (2021) Reference frame induced symmetry breaking on holographic screens. Symmetry 13, 408. https://doi.org/10.3390/sym13030408

Fields, C., Glazebrook, J. F. and Marcianò, A. (2022a) Sequential measurements, topological quantum field theories, and topological quantum neural networks. Fortschr. Physik 70, 202200104. https://doi.org/10.1002/prop.202200104

Fields, C.; Glazebrook, J. F.; Marcianò, A. (2022b) The physical meaning of the holographic principle. *Quanta* 11, 72–96.

Fields, C.; Glazebrook, J. F.; Marcianò, A. (2023) Communication protocols and quantum error-correcting codes from the perspective of topological quantum field theory. Preprint arxiv:2303.16461.

Fields, C. and Levin, M. (2019) Somatic multicellularity as a satisficing solution to the prediction-error minimization problem. Commun. Integr. Biol. 12, 119-132. https://doi.org/10.1080/19420889.2019.1643666

Fields, C. and Levin, M. (2020a) Scale-free biology: Integrating evolutionary and developmental thinking. BioEssays 2020, 1900228. https://doi.org/10.1002/bies.201900228

Fields, C. and Levin, M. (2020b) Does evolution have a target morphology? Organisms 4, 57-76. https://doi.org/10.13133/2532-5876/16961

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

Fields, C. and Levin, M. (2021) Metabolic limits on classical information processing by biological cells. BioSystems 209, 104513. https://doi.org/10.1016/j.biosystems.2021.104513

Fields, C. and Marcianò, A. (2019) Sharing nonfungible information requires shared nonfungible information. Quant. Rep. 1, 252-259. https://doi.org/10.3390/quantum1020022

Fields, C. and Marcianò, A. (2020) Holographic screens are classical information channels. Quant. Rep. 2, 326-336. https://doi.org/10.3390/quantum2020022

Friston, K. J. (2005) A theory of cortical responses. Philos. Trans. R. Soc. Lond. B, Biol. Sci. 360, 815-836.

Friston, K. J. (2010) The free-energy principle: A unified brain theory? Nature Rev. Neurosci. 11, 127- 138.

Friston, K. J. (2013) Life as we know it. J. R. Soc. Interface 10, 20130475.

Friston, K. J. (2019) A free energy principle for a particular physics. Preprint arXiv:1906.10184 [q- bio.NC].

Friston, K., Da Costa, L., Sakthivadivel, D. A. R., Heins, C., Pavliotis, G. A., Ramstead, M. and Parr, T. (2022) Path integrals, particular kinds, and strange things. Preprint arxiv:2210.12761.

Friston, K. J., FitzGerald, T., Rigoli, F., Schwartenbeck, P. and Pezzulo, G. (2017) Active inference: A process theory. Neural Comput. 29, 1-49.

Friston, K. J., Kilner, J. and Harrison, L. (2006) A free energy principle for the brain. J. Physiol. (Paris) 100, 70-87.

Friston, K., Parr, T., Yufik, Y., Sajid, N., Price, C. J. and Holmes, E. (2020) Generative models, linguistic communication and active inference. Neurosci. Biobehav. Rev. 118, 42-64.

Froese, T. and Taguchi, S. (2019) The problem of meaning in AI and robotics: Still with us after all these years. Philosophies 4, 14. https://doi.org/10.3390/philosophies4020014

Fuchs, C.A. (2003) Quantum mechanics as quantum information, mostly. J. Mod. Opt. 50, 987-1023.

Fuchs, C., 2010. QBism, the perimeter of quantum Bayesianism. Preprint arxiv:1003.5209.

Fuchs, C. and Schack, R. (2013) Quantum Bayesian coherence. Rev. Mod. Phys. 85, 1693-1715.

Gidon, A., Zolnik, T. A., Fidzinski, P., Bolduan, F., Papoutsi, A., Poirazi, P., Holtkamp, M., Vida, I. and Larkum, M. E. (2020) Dendritic action potentials and computation in human layer 2/3 cortical neurons. Science 367, 83-87.

Gilbert, S. F. (2014a) Symbiosis as the way of eukaryotic life: The dependent co-origination of the body. J. Biosci. 39, 201-209.

Gilbert, S. F. (2014b) A holobiont birth narrative: the epigenetic transmission of the human microbiome. Front. Genet. 5, 282.

Guerrero, R., Margulis, L. and Berlanga, M. (2013) Symbiogenesis: The holobiont as a unit of evolution. Int. Microbiol. 16, 133-143.

Jumper, J., R. Evans, A. Pritzel et al. (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596, 583-589.

Li, Y., D. Choi, J. Chung et al. (2022) Competition-level code generation with AlphaCode. Science 378, 1092-1097.

Hoel, E. P. (2017) When the map is better than the territory. Entropy 19, 188.

Hoel, E. P., Albantakis, L., Tononi, G. (2013) Quantifying causal emergence shows that macro can beat micro. Proc. Natl. Acad. Sci. USA 110, 19790-19795.

Horsman, C., Stepney, S., Wagner, R. C. and Kendon, V. (2014) When does a physical system compute? Proc. R. Soc. A 470, 20140182.

Ingber, D. E. (2008) Can cancer be reversed by engineering the tumor microenvironment? Semin. Cancer Biol. 18, 356-364.

Kirchhoff, M., Parr, T., Palacios, E., Friston, K. and Kiverstein, J. (2018) The Markov blankets of life: autonomy, active inference and the free energy principle. J. R. Soc. Interface 15, 2017.0792. https://doi.org/10.1098/rsif.2017.0792

Kondo, S. and Miura, T. (2010) Reaction-diffusion model as a framework for understanding biological pattern formation. Science 329, 1616-1620.

Kramer, B. A., del Castillo, J. S. and Pelkmans, L. (2022) Multimodal perception links cellular state to decision-making in single cells. Science 377, 642-648. https://doi.org/10.1126/science.abf4062

Kriegman, S., Blackiston, D., Levin, M. and Bongard, J. (2021) Kinematic self-replication in reconfigurable organisms. Proc. Natl. Acad. Sci. USA 118, e2112672118. https://doi.org/10.1073/pnas.2112672118

Kuchling, F., Fields, C. and Levin, M. (2022) Metacognition as a consequence of competing evolutionary time scales. Entropy 24, 601.

Kuchling, F., Friston, K., Georgiev, G. and Levin, M. (2020) Morphogenesis as Bayesian inference: A variational approach to pattern formation and control in complex biological systems. Phys. Life Rev. 33, 88-108.

Kuehner, J. N. and Brow, D. A. (2006) Quantitative analysis of in vivo initiator selection by yeast RNA Polymerase II supports a scanning model. J. Biol. Chem. 281, 14119-14128. https://doi.org/10.1074/jbc.M601937200

Landauer, R. (1961) Irreversibility and heat generation in the computing process. IBM J. Res. Devel. 5, 183-195.

Landauer, R. (1999) Information is a physical entity. Physica A 263, 63-67.

Levin, M. (2011) The wisdom of the body: Future techniques and approaches to morphogenetic fields in regenerative medicine, developmental biology and cancer. Regen. Med. 6, 667-673.

Levin, M. (2019) The computational boundary of a "self": Developmental bioelectricity drives multicellularity and scale-free cognition. Front. Psychol. 10, 2688.

Levin, M. (2021) Life, death, and self: Fundamental questions of primitive cognition viewed through the lens of body plasticity and synthetic organisms. Biochem. Biophys. Res. Comm. 564, 114-133. https://doi.org/10.1016/j.bbrc.2020.10.077

Levin, M. (2022) Technological approach to mind everywhere: An experimentally-grounded framework for understanding diverse bodies and minds. Front. Syst. Neurosci. 16, 768201. https://doi.org/10.3389/fnsys.2022.768201

Lobo, D., Solano, M., Bubenik, G. A. and Levin, M. (2014) A linear-encoding model explains the variability of the target morphology in regeneration. J. R. Soc. Interface 11, 20130918.

Lovelock, J. E. and Margulis, L. (1974) Atmospheric homeostasis by and for the biosphere: The Gaia hypothesis. Tellus 26, 2-10.

Lyon, P (2015) The cognitive cell: Bacterial behavior reconsidered. Front. Microbiol. 6, 264.

Lyon, P. (2020) Of what is "minimal cognition" the half-baked version? Adapt. Behav. 28: 407-428.

Maturana, H. R. and Varela, F. J. (1980) Autopoesis and Cognition: The Realization of the Living. Dordrecht, D. Reidel.

McMillan, P., Oudin, M. J., Levin, M. and Payne, S. L. (2021) Beyond neurons: Long distance communication in development and cancer. Front. Cell Dev. Biol. 9, 739024.

Mermin, N.D. (2018) Making better sense of quantum mechanics. Rep. Prog. Phys. 82, 012002.

Michod, R. E. (1999) Darwinian Dynamics. Princeton University Press, Princeton, NJ.

Monod, J. (1972) Chance and Necessity. Random House, New York.

Morelli, L. G., Uriu, K., Ares, S. and Oates, A. C. (2012) Computational approaches to developmental patterning. Science 336, 187-191.

Nielsen, M. A. and Chuang, I. L. (2000) Quantum Computation and Quantum Information. Cambridge, UK: Cambridge University Press.

Parrondo, J. M. R., Horowitz, J. M. and Sagawa, T. (2015) Thermodynamics of information. Nat. Phys. 11, 131-193. https://doi.org/10.1038/NPHYS3230

Pattee, H. H. (1982) Cell psychology. Cognit. Brain Theory 5, 325-341.

Pearl, J. (1988) Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Mateo CA.

Pegg, D. T., Barnett, S. M. and Jeffers, J. (2002) Quantum theory of preparation and measurement. J. Mod. Optics 49, 913-924. https://doi.org/10.1080/09500340110109412

Pezzulo, G. and Levin, M. (2016) Top-down models in biology: Explanation and control of complex living systems above the molecular level. J. R. Soc. Interface 13, 20160555.https://doi.org/10.1098/rsif.2016.0555

Pinet, K. and McLaughlin, K. A. (2019) Mechanisms of physiological tissue remodeling in animals: Manipulating tissue, organ, and organism morphology. Dev. Biol. 451, 134-145.

Polanyi, M. (1968) Life's irreducible structure. Live mechanisms and information in DNA are boundary conditions with a sequence of boundaries above them. Science 160, 1308-1312. https://doi.org/10.1126/science.160.3834.1308

Raja, V., Valluri, D., Baggs, E., Chemero, A. and Anderson, M. L. (2021) The Markov blanket trick: On the scope of the free energy principle and active inference. Phys. Life Rev. 39, 49-72.

Ramstead, M. J., Constant, A., Badcock, P. B. and Friston, K. J. (2019) Variational ecology and the physics of sentient systems. Phys. Life Rev. 31, 188-205.

Ramstead, M. J., Sakthivadivel, D. A. R., Heins, C., Koudahl, M., Millidge, B., Da Costa. L., Klein, B., and Friston, K. J. (2022) On Bayesian mechanics: A physics of and by beliefs. Preprint arXiv:2205.11543 [cond-mat.stat-mech].

Rosen, R. (1986) On information and complexity. In: J. L. Casti, A. and Karlqvist, A. (Eds.) Complexity, Language, and Life: Mathematical Approaches. Springer, Berlin, pp. 174-196.

Samek, W., G. Montavon, G., Lapuschkin, S., Anders, C. J. and Müller, K.-R. (2021) Explaining deep neural networks and beyond: A review of methods and applications. Proc. IEEE 109, 247-278.

Sakthivadivel, D. A. R. 2022 A constraint geometry for inference and integration. Preprint arXiv:2203.08119.

Sakthivadivel, D. A. R. 2022 Towards a geometry and analysis for Bayesian mechanics. Preprint arXiv:2204.11900.

Sakthivadivel, D.A.R. 2022 Weak Markov blankets in high-dimensional, sparsely-coupled random dynamical systems. Preprint arXiv:2207.07620.

Smith, R., Babcock, P. and Friston, K. J. (2020) Recent advances in the application of predictive coding and active inference models within clinical neuroscience. Psychiatry Clin. Neurosci. 75, 3-13.

Smith, J. E. and Nair, R. (2005) The architecture of virtual machines. IEEE Computer 38(5), 32-38. https://doi.org/10.1109/MC.2005.173

Stewart, J. (1996) Cognition = Life: Implications for higher-level cognition. Behav. Process. 35, 311- 326.

Strassmann, J. E. and Queller, D. C. (2010) The social organism: Congresses, parties and committees. Evolution 64, 605-616. https://doi.org/10.1111/j.1558-5646.2009.00929.x

Szathmáry, E. and Maynard Smith, J .(1995) The major evolutionary transitions. Nature 374, 227-232.

Tegmark, M. (2012) How unitary cosmology generalizes thermodynamics and solves the inflationary entropy problem. Phys. Rev. D 85, 123517.

Tintori, S. C., Nishimura, E. O., Golden, P., Lieb, J. D. and Goldstein, B. (2016) A transcriptional lineage of the early C. elegans embryo. Devel. Cell 38, 430-444. https://doi.org/10.1016/j.devcel.2016.07.025

Toyabe, S., Sagawa, T., Ueda, M., Muneyuki, E. and Sano, M. (2010) Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality. Nat. Phys. 6, 988-992. https://doi.org/10.1038/nphys1821

Turing, A. (1953) The chemical basis of morphogenesis. Phil. Trans. R. Soc. London 237, 37-72.

Vandenberg, L. N., Adams, D. S. and Levin, M. (2012) Normalized shape and location of perturbed craniofacial structures in the Xenopus tadpole reveal an innate ability to achieve correct morphology. Dev. Dyn. 241, 863-878.

Wang, M., Hu, Q., Lv, T., Wang, Y., Lan, Q., Xiang, R. et al. (2022) High-resolution 3D spatiotemporal transcriptomic maps of developing Drosophila embryos and larvae. Devel. Cell 57, 1271-1283. https://doi.org/10.1016/j.devcel.2022.04.006

Wheeler, J. H. (1989) Information, physics, quantum: The search for links. In: Zurek, W. (Ed.), Complexity, Entropy, and the Physics of Information. CRC Press, Boca Raton, FL, pp. 3-28.

Zurek, W. H. (2003) Decoherence, einselection, and the quantum origins of the classical. Rev. Mod. Phys. 75, 715-775.

Zurek, W.H. (2009) Quantum Darwinism. Nat. Phys. 5, 181-188.