

# A Multi-scale View of the Emergent Complexity of Life: A Free-energy Proposal

## Abstract

We review some of the main implications of the free-energy principle (FEP) for the study of the self-organization of living systems – and how the FEP can help us to understand (and model) biotic self-organization across the many temporal and spatial scales over which life exists. In order to maintain its integrity as a bounded system, any biological system - from single cells to complex organisms and societies - has to limit the disorder or dispersion (i.e., the long-run entropy) of its constituent states. We review how this can be achieved by living systems that minimize their variational free energy. Variational free energy is an information theoretic construct, originally introduced into theoretical neuroscience and biology to explain perception, action, and learning. It has since been extended to explain the evolution, development, form, and function of entire organisms, providing a principled model of biotic self-organization and autopoiesis. It has provided insights into biological systems across spatiotemporal scales, ranging from microscales (e.g., sub- and multicellular dynamics), to intermediate scales (e.g., groups of interacting animals and culture), through to macroscale phenomena (the evolution of entire species). A crucial corollary of the FEP is that an organism just is (i.e., embodies or entails) an implicit model of its environment. As such, organisms come to embody causal relationships of their ecological niche, which, in turn, is influenced by their resulting behaviors. Crucially, free-energy minimization can be shown to be equivalent to the maximization of Bayesian model evidence. This allows us to cast evolution (i.e. natural selection) in terms of Bayesian model selection, providing a robust theoretical account of how organisms come to match or accommodate the spatiotemporal complexity of their surrounding niche. In line with the theme of this volume; namely, biological complexity and self-organization, this chapter will examine a variational approach to self-organization across multiple dynamical scales.

**Keywords:** Free-energy principle; Active inference; Self-organization; Markov blanket; Niche Construction; Variational neuroethology

**Authors:** Hesp, Ramstead, Constant, Badcock, Kirchhoff, & Friston

# Introduction

The emergence of life - and biological self-organization - is a fascinating topic for many working within the life sciences, as well as to laypersons and scholars outside biology. We review an integrative account of the self-organization of life across temporal and spatial scales, based on the free-energy principle<sup>1</sup> (FEP, for short). Any view on biological self-organization must explain how organisms remain alive; that is, resist systematic dispersion and entropic decay. Organisms need to retain a grasp on their own environment in order to maintain their integrity; i.e., structure and function. For example, bacteria have implicit expectations about the temperature range in which their metabolism fares best (resulting in behavior called thermotropism). In this way, they resist the natural tendency towards decay or disorder. More generally, organisms embody expectations that they need to ensure are brought about through adaptive action.

Tropism in bacteria is an example of how organisms do not just passively predict their sensory states, but act on their environment to realize their own expectations (e.g., concerning their preferred temperature). In other words, an organism's behavior can be cast in terms of self-fulfilling prophecies; what we call *active inference* (Friston, Daunizeau, & Kiebel 2009). Organisms need implicit beliefs about the outer world (like the direction of a heat source) to bring about an adaptive action (moving away from the heat). Yet, they never have direct access to the outer world; only to what impinges upon their sensory receptors. Conversely, the outer world is influenced by the actions of the organism, but not by its inner states.

Thus, active inference forms a circle, from the inner world of the organism to its actions on the outer world, which feeds back to the organism through sensory stimulation. What makes this circularity virtuous rather than vicious is the information-theoretic concept of *variational free energy* (Friston 2010; 2013). Variational free energy is a measure of *the difference between what the organism senses and what it expects to sense*. Technically, variational free energy is an upper bound on 'self information', 'surprisal', or simply 'surprise', which reflects how surprising (or improbable) the current state of the world is for the organism (including its internal states). Although surprise itself cannot be evaluated explicitly by the organism, variational free energy can be; because it depends only on probabilistic beliefs about the world 'out there', which are encoded by the state of the organism. Thus, variational free energy is a proxy for surprise.

The time average of surprise (i.e., self information) is informational entropy. This entropy is a measure of uncertainty, which means that free energy effectively places an upper limit on the entropy of organism's sensory exchanges with the world and – if it acts in a way that minimizes expected free energy – uncertainty about its lived world. Free-energy minimization can be pursued in many ways; it has been suggested that it is an explanatory principle flexible enough to incorporate many (and possibly all) phenomena studied under the rubric of cognition (Badcock 2012; Clark 2015; Friston 2010; Hohwy 2013).

---

<sup>1</sup> The term free energy has been used with and without hyphenation in the literature. Throughout this chapter, we write "free energy" when used as a noun (e.g., organisms minimize free energy), and "free-energy" when used as an adjective (e.g., free-energy principle and free-energy minimization).

Crucially, because minimizing free energy places an upper bound on surprisal, it is equivalent to placing a lower bound on Bayesian model evidence (i.e., negative surprisal) for an implicit model (i.e., the organism) that produces expectations about sensory data. As such, free-energy minimization corresponds to a form of *variational* or *approximate Bayesian inference*, widely employed in machine-learning and statistics (Friston, 2010; Kirchoff et al., 2018; Ramstead, Badcock, & Friston, 2017). This recurrent, incremental process of optimization is by its nature approximate because organisms (and machines) do not have direct access to the outer world (in a statistical sense). Organisms themselves *are* the implicit model for which they gather evidence, resulting in the interpretation that they produce evidence for their own existence - they are effectively self-evidencing (Hohwy 2016). This self-referential recurrence is central to active inference, in which all of life engages perpetually. We can therefore use approximate Bayesian inference and associated (implicit) probabilistic beliefs to characterize the interactions of an organism with its local niche.

If biological systems did not minimize free energy efficiently, the disorder or entropy of their sensory states would not be sufficiently bounded and diverge, leading to disintegration and death (in accord with the fluctuation theorem that generalizes the second law of thermodynamics to open systems). Therefore, biological systems must minimize free energy. More generally, this line of reasoning suggests that any complex adaptive (sub)system that underwrites its own existence will minimize free energy and therefore engage in active inference with respect to its surrounding environment (Friston 2010; 2013). Indeed, later on we illustrate how random dynamical systems can give rise to such inferential behaviors (Section 3.1).

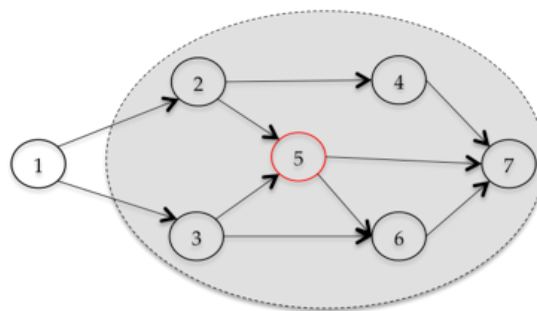
**Special care needs to be taken when relating the information-theoretic constructs that are employed in the variational free energy formulation to thermodynamic constructs such as Gibbs entropy and Gibbs free energy. This step is important if the FEP is to act as an integrative scientific framework that leverages, and connects to, the physical sciences in the study of biological self-organization. We emphasize that variational free energy is conceptually distinct from thermodynamic free energy. The fact that both quantities share the same label (i.e., “free energy”) derives from their analogous mathematical definitions. Otherwise, the relationship between the two quantities is non-trivial and much of the work relating them remains to be done (Ramstead et al., 2018b; see, e.g., Sengupta, Stemmler, & Friston 2013, for an account of this connection based on neuronal processing efficiency). The same holds for information-theoretic entropy and thermodynamic entropy, although these two constructs are more closely and straightforwardly related (e.g., through Boltzmann’s famous entropy formula). The difficulty in relating these concepts stems largely from the fact that the FEP operates in a different regime from that usually considered under statistical physics. The FEP is formulated appropriately for the study of biological self-organization, since it pertains to systems at non-equilibrium steady state (NESS); whereas statistical mechanics focuses primarily on equilibrium (or near-equilibrium) states that allow for robust descriptions of physical systems in a particular equilibrated state. Having said this, the FEP and thermodynamics are internally consistent in the sense that thermodynamics – particularly stochastic thermodynamics (Ao, 2008; Seifert, 2012) – can be regarded as a special case of the FEP when certain conditions are met (Friston & Ao, 2012).**

With the above caveat in mind, we devote this chapter to reviewing the implications of the FEP for explaining the adaptive self-organization of living systems across different spatiotemporal scales, ranging from microscales (e.g., cells) to intermediate scales (e.g., learning processes of animals), and eventually to the evolutionary macroscale (i.e., the emergence of entire species). We suggest that once the FEP is extended to these different scales of self-organization, these processes, which might appear miraculous, are not really as ‘surprising’ as one might have thought. The events that take place within the boundary of a living organism arise from the very existence of that boundary (called the Markov blanket, as explained below), the emergence of which is itself nearly inevitable in a physically lawful world like ours.

The structure of the chapter is as follows. In Section 1, we introduce the concept of a Markov blanket and its relation to free-energy minimization and active inference. In Section 2, we generalize active inference across spatiotemporal scales, to formulate a multi-scale interpretative framework for biological self-organization. In Section 3, we examine some examples of active inference at the sub- and multi-cellular microscale, notably demonstrating how active inference: (i) emerges directly from a primordial soup; (ii) channels dendritic self-organization of single neurons; and (iii) enables the collective organization of many cells into entire organs. In Section 4, we turn to the organismic level, where we consider: (i) the hierarchical brain; (ii) communication and dialogue through active inference; and (iii) cultural affordances and collective active inference. In Section 5, we consider the species macroscale. We first discuss how biological evolution can be viewed as a form of active inference over the order parameters of the lower levels treated in Sections 3-4. Finally, we focus on niche construction, and examine its role throughout both development and evolution to describe how species build their own eco-niche.

# 1. Markov Blankets and Active Inference

A key aspect of living systems is that they function adaptively by means of their own self-perpetuating, self-organizing boundaries (Varela, Maturana, & Uribe 1974). Adaptive self-organization enables a living system to establish and maintain a boundary that separates its internal states from the states comprising its external milieu (Barandiaran & Moreno 2008), which in turn allows for active inference. This type of boundary can be viewed as a Markov blanket. Pearl (1988) introduced the notion of a Markov blanket to denote a set of epistemological properties specific to Bayesian networks (Figure 1). The Markov blanket is cast as the smallest set of nodes that renders an enclosed node conditionally independent of all others. The central point is that the behavior of the enclosed node can be predicted by knowing only the states of the nodes that constitute its Markov blanket. Nodes outside the Markov blanket provide no additional information. Conversely, when predicting the behavior of the nodes outside the Markov blanket, the enclosed node provides no additional information beyond that provided by the Markov blanket itself.



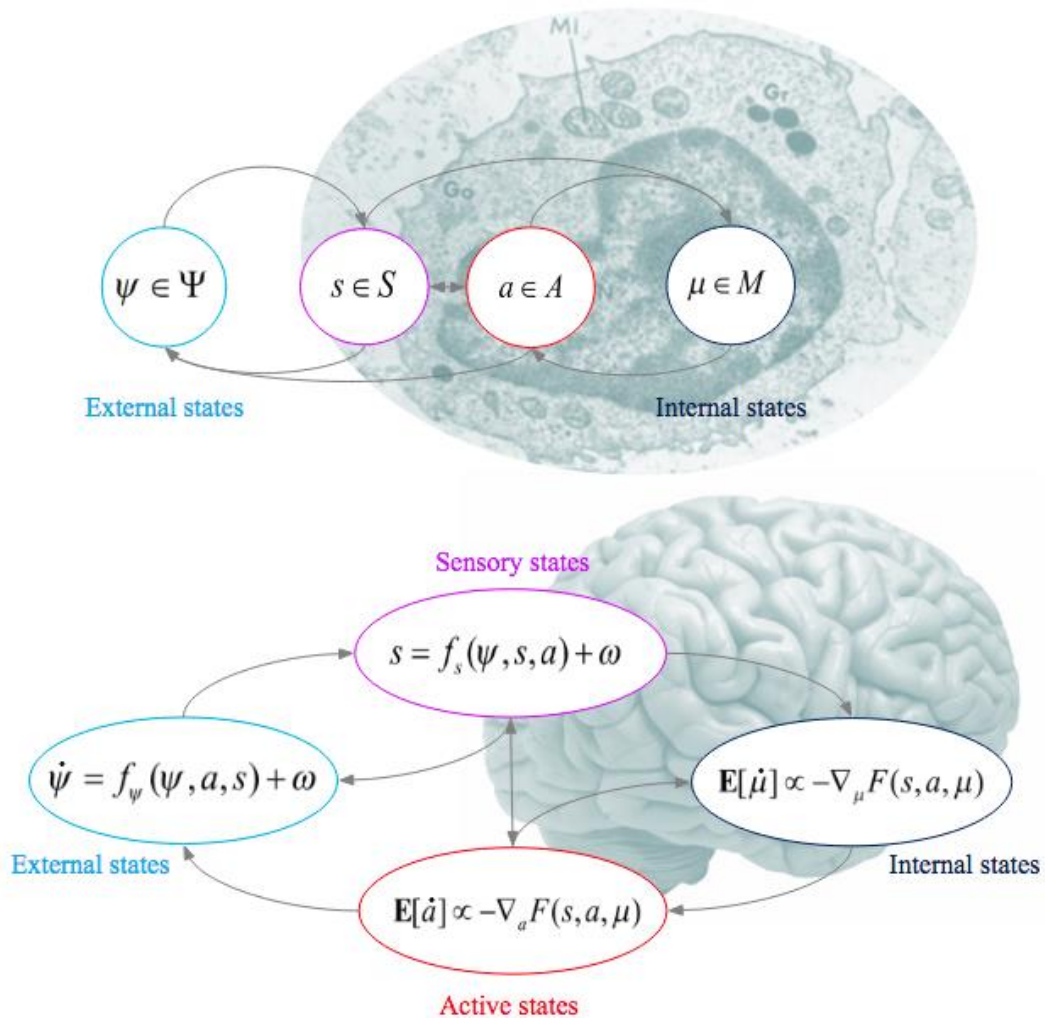
**Figure 1.** A graphical depiction of a Markov blanket with full conditionals. Nodes represent random variables and arrows or edges represent conditional dependencies. In this figure, the Markov blanket for node (Sun, Gomez, #252, & Schmidhuber) is the union of its *parents* {2,3}, the *children* or direct successors of (Sun et al.), which are {6,7}, and the *parents' children* {4}. Hence, (Sun et al.) = {6,7}  $\cup$  {2,3}  $\cup$  {4} = {2,3,4,6,7}. The union of (Sun et al.) does not include {1}. This implies that {1} and (Sun et al.) are conditionally independent given {2,3,4,6,7}, and shows that once the union of (Sun et al.) is given, the probability of (Sun et al.) will not be affected by the probability of {1}. Formally, (Sun et al.) is conditionally independent of {1} given {2,3,4,6,7}, if  $P((\text{Sun et al.})|\{1\}, \{2,3,4,6,7\}) = P((\text{Sun et al.})|\{2,3,4,6,7\})$ . This means that once all the neighboring variables for (Sun et al.) are known, knowing the state of {1} provides no additional information about the state of (Sun et al.). It is this kind of statistical neighborhood for (Sun et al.) that is called a Markov blanket (Pearl 1988). This figure is from (Kirchhoff et al., 2018); adapted from Murphy 2012, p. 329).

The notion of a Markov blanket, and the independencies between states it induces, can be directly applied to biological systems (Friston, 2013; Palacios et al., 2017). For example, the interior of a cell can be related to the internal states of the cell (e.g., cell metabolism), the extracellular environment to its external states, and the cell boundary to the Markov blanket that couples intracellular and extracellular states to one another. The states that constitute

the Markov blanket can be further partitioned into sensory and active states. As such, the presence of a Markov blanket implies a partitioning of states into external, sensory, active and internal states (see Figure 2; Friston et al. 2015).

Figure 2 highlights the partitioning rule governing the Markov blanket formalism; namely, that hidden external states influence sensory states, which influence, but are not themselves influenced by, internal states. Conversely, internal states influence active states, which influence, but are not themselves influenced by, external states. This formulation relies on the statistical dependencies between the states comprising a biological system - internal states and their Markov blanket - and the kind of independencies induced between internal and external states. Importantly, this formulation echoes key themes of dynamical coupling between the organism and its environment in enactive and embodied approaches to biology and cognition (Engel, Friston, & Kragic, 2016; Noë, 2004; Thompson, 2007; Varela, Thompson, & Rosch, 2017).

The dependencies established by a Markov blanket induce *active inference*, which rests on the principle that adaptive action reduces uncertainty or surprise about the causes of sensory data (Mirza et al. 2016). The statistical properties of Markov blankets result in emerging (self-organizing) processes that optimize Bayesian model evidence, such that it becomes possible to associate the internal states of a system with a model of the external states (Friston et al. 2015; Kirchhoff et al., 2018). Action, which is induced by the generation of inferences via internal states, drives an organism toward a free-energy minimum (Parr & Friston 2018). We will develop this point in further detail as we move through the various sections of our review.



**Figure 2.** These two illustrations highlight the dependencies between states induced by the presence of the Markov blanket of a cell (top) and the brain (bottom). Internal states (black) are connected to the external states (blue) through the sensory (magenta) and active (red) states. (Figure taken from Friston (2013; Figure 1).

## 2. Active Inference at Multiple Scales

The variational approach has recently been extended to explicitly address living systems across spatial and temporal scales ([Kirchhoff et al., 2018](#); [Ramstead et al., 2017](#)), relying heavily on the concept of a Markov blanket introduced in Section 1. Any (ergodic) system that exists must, in virtue of existing, be enshrouded by a Markov blanket that maintains it. This holds for the component states of any Markov blanketed system as well. In principle, we can describe the universe of biological systems as Markov blankets and their internal states, which are themselves composed of Markov blankets and their internal states. This formalism can be reiterated all the way up, and all the way down; i.e., across the manifold nested scales of organization at which biological systems exist, including their eco-niche. In this way, biotic systems (i.e., single cells, organisms, social and cultural groups) can be described as a (high-dimensional) phase space that is induced by a hierarchy of Markov blankets. This view of living systems has been labeled *variational neuroethology* ([Ramstead et al., 2017](#)). As humans, we are a prime example: our brains, sensory organs, and muscles are themselves composed of countless cells, each possessing their own Markov boundary.

This multiscale extension of the Markov blanket formalism involves the notion of a *scale space*. Scale spaces enable us to carve out different structures at different spatial and temporal scales, and to flag which kinds of systems are relevant to our investigations at those scales. In this context, scale spaces are useful because they allow us to model the dynamics of integrated nested systems; that is, how systems at one scale produce or entail the composite system at a higher level. Moving up the hierarchy of Markov blankets entails an increase in spatial and temporal scales. Any system that can be distinguished from its environment (and thus, possesses a Markov blanket) can take part in a dynamical interaction that produces a Markov blanket at a higher level of organization ([Palacios et al., 2017](#)).

By way of illustration, consider an ensemble of cells, each bounded by their respective plasmalemmas. We can mathematically model the self-organization of the cellular ensemble by appealing to the dynamic interactions between their sensory and active states, shaped by their collective effort to minimize free energy. Exchanges at one scale (e.g., the scale of cellular interactions) have a sparsity structure that, in turn, can induce a Markov blanket at the scale above. For example, some group of cells in that ensemble could be epithelial cells that, in turn, constitute the boundary of an entire organ. Conversely, within the cell, the various organelles have their own Markov blanket. Despite the difference in scale, the dynamics involved have a formally identical statistical structure; namely, that prescribed by the Markov blanket formalism.

The hierarchical nesting of Markov blankets provides a vantage point from which to model the self-organization of biological systems across spatial and temporal scales. Crucially, it also provides a principled explanation of how each level contextualizes (that is, constrains) ongoing dynamics at other scales. The very same variational, entropy-bounding dynamics are operative at each scale, and provide an integrative dynamics for the entire system. Free-energy minimization unifies these various scales and allows them to be evaluated simultaneously. In the following sections, we will first address the emergence of



the Markov blanket and then proceed to explore the application of the free-energy principle to the various scales at which life exists.

## 3. Microscale: Sub- and Multicellular Self-organization

### 3.1 Emergence of Markov blankets and active inference in a primordial soup

A complete treatment of the origins of life would have to address the emergence of prokaryotic cells and their capability to produce descendants that carry their (epi)genetic inheritance. As noted before, the structure and function of the cell is a prime example of how Markov blankets induce active inference. In line with these insights, we choose to first address how random dynamical systems can give rise to sub-systems that maintain themselves through active inference (Friston, 2013). This is a crucial step, because once such a “primal Markov blanket” is established, the sub-system becomes self-sustaining and, hence, susceptible to innovations and organization into larger composite systems. For example, it is thought that some of the organelles within eukaryotic cells used to be prokaryotic cells themselves (i.e., mitochondria and chloroplasts). Although this is far from a full account of life as we know it, we can use abstract representations of dynamical processes to illustrate some simple but fundamental aspects of adaptive self-organization. These processes may serve as a metaphor for dynamical interactions across various levels of biological self-organization.

The following theorem will serve as a guideline in what follows: *if a random dynamical system is ergodic and has a Markov blanket, it actively maintains its own structure and dynamics* (i.e., autopoiesis; Friston 2013). Ergodicity is a key concept, which formally means that the average of any measurable function within the system converges over time. This definition implies that a limited number of states are being revisited, because not all functions would converge for an infinite number of possible states. By virtue of ergodicity, the average proportion of time a certain state is occupied (within a sufficiently large window) is equivalent to the probability of the system being in that state when observed at random. In other words, an ergodic random dynamical system is tractable in terms of probabilities, which is crucial for any type of inference. Ergodicity is readily identified as a key property of biological systems. For example, neurons switch between their resting, firing, and refractory states.

Friston (2013) provided a proof of principle of this simple but fundamental property of living systems. He modelled a “primordial soup” that exhibited the type of behavior described in the theorem presented above. These simulations consisted of a collection of dynamical subsystems, which can be likened to macromolecules. Each of these macromolecules could reside in a number of possible structural and functional states, and was coupled by these states with other nearby macromolecules. The type of dynamics employed in these simulations is similar to those in the wealth of literature on pattern formation in dissipative systems; e.g., turbulence in hydrodynamics (e.g., Manneville, 1995). In the context of Friston’s simulations, structural states represented the locations and motions of these macromolecules, while functional states represented their electrochemical states. Through electrochemical interactions, functional states can influence the location and velocities (structural states) of nearby molecules, as well as the electrochemical states of those

molecules. The intention of this exercise is not to analyze the precise patterns that emerge from these interactions, but rather to demonstrate that a basic form of active inference can emerge from a “primordial soup”.

While each of the sub-systems themselves only has a limited number of possible functional states (i.e., they are locally ergodic), the simulations also exhibited emergent ergodic behavior for the system as a whole. Initially, macromolecules pushed each other away; after a few cycles, they tended to clot together, forming a stable dense clump. Short-distance interactions led to a pattern in which macromolecules were passed around until they only gently pushed and pulled on each other most of the time, with occasional bursts of movement. The collective motion and electrochemical states of this dense emerging clump could be characterized as a “restless soup”, as shown in Figure 3.

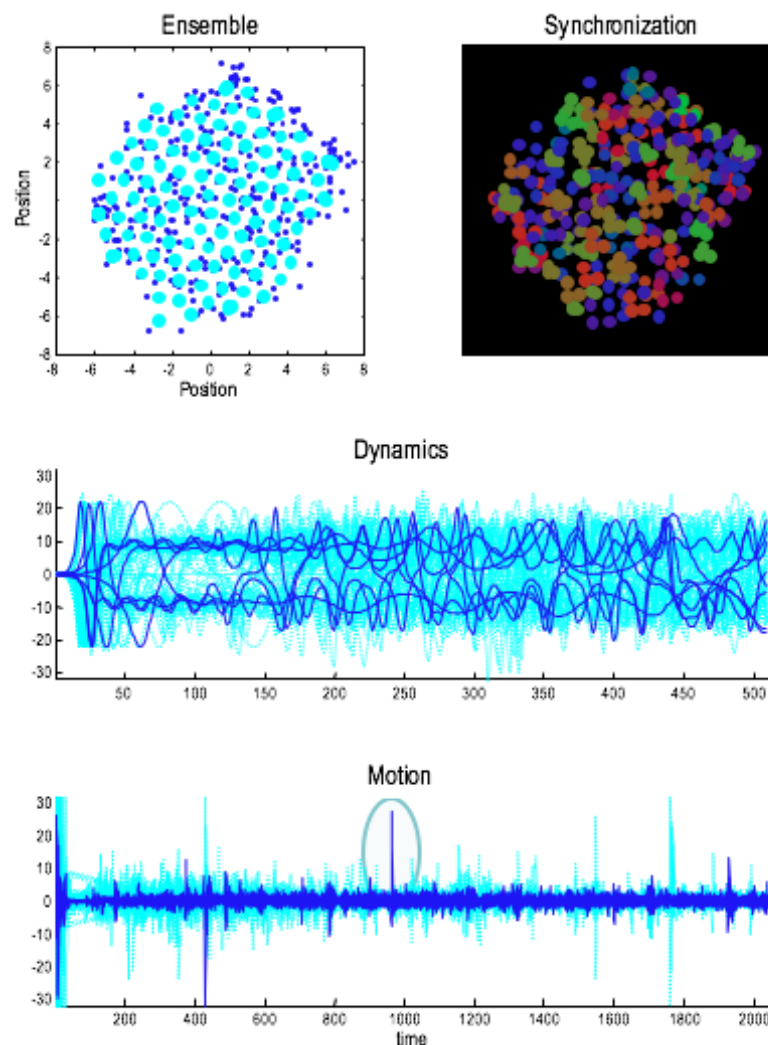
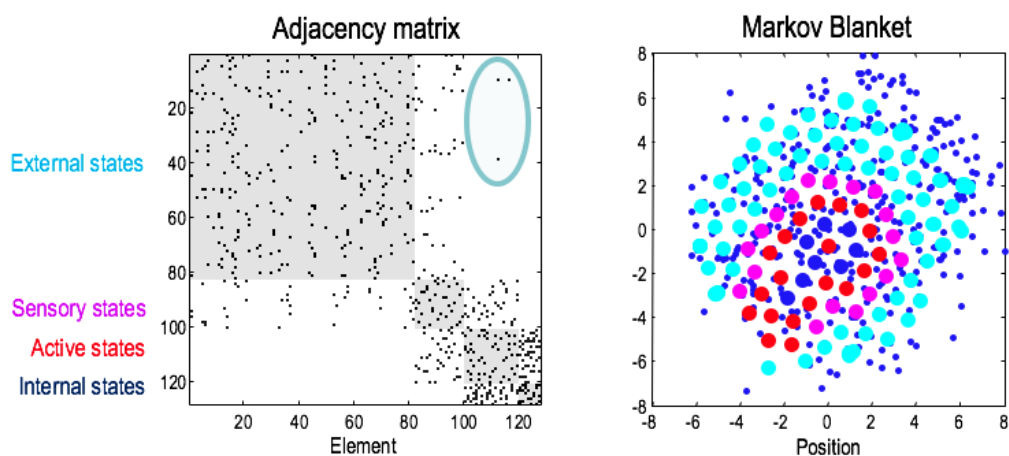


Figure 3. Reproduced from Friston (2013; Figure 1), this figure shows the structure and temporal dynamics of the simulated primordial soup. Panel a(i) illustrates the spatial position (large cyan dot) and functional states (three dark blue dots) for each of the 128 subsystems, after the states have converged on their global (random dynamical) attractor. Panel a(ii) shows the same snapshot of time with the three functional states coded by color,

illustrating the synchronization of electrochemical states across the clump. Panels b and c show, respectively, the functional states and motion as a function of time (in seconds, processor time). internal states are shown in blue and external states in cyan. The circle in panel c indicates one of the occasional bursts of motion due to the non-linear dynamics within the clump of macromolecules. See Friston (2013) for technical details.

Is there any active inference evident in this synthetic mess? Given that the global attractor state of the system as well as the subsystems themselves are ergodic, we can characterize their behaviors in probabilistic terms. We can then use the coupling between the states of these macromolecules to disentangle their spheres of influence. Based on this information, we can identify the Markov blanket (if present), and the states enclosed by it. Friston (2013) found that amidst the densest region of the “soup” were a number of macromolecules that were very tightly coupled to one another, and whose states were completely hidden from those residing on the outer edges of the system. Figure 4 shows the macromolecules representing internal states (dark blue) and those representing the Markov blanket as the sensory (magenta) and active (red) states. The active macromolecules, which allow the internal states to affect the outer world indirectly, lie within the sensory subsystems that are exposed to the outer world. Interestingly, biological cells have a somewhat similar configuration, with an (active) cytoskeleton surrounded by (sensory) epithelia or receptors.



**Figure 4.** This figure shows the emergence of the Markov blanket from the primordial soup after the global attractor state was reached. The left panel shows the coupling between the 128 macromolecules over 256 seconds (adjacency matrix), ordered according to the internal (blue), active (red), sensory (purple), and external hidden (cyan) subsystems. The circle indicates instances of active subsystems influencing external states (owing to the periodic bursts of motion) without the external states influencing the active states. The right panel shows the spatial organization of this partition. Reproduced from Friston (2013; Figure 2).

Crucially, a minimalistic form of perception was also identified within the clump of macromolecules. Although particles in the interior were entirely insulated from the outer

world, their functional states were shown to have predictive value for the motion of the macromolecules *outside* the clump. In a self-organized fashion, these mindless, simplistic “representations” of macromolecules appeared to be producing implicit inferences about the world outside their synthetic bubble.

Friston also showed that the implicit inferences – driven by the (sensorial) dynamics of the inner environment of the clump – directed the active states to maintain its structure. In this way, the clump of macromolecules essentially anticipated future perturbations induced by the outer world, and acted on these expectations: a basic form of active inference.

We can now return to the theorem introduced above. Does the emergent clump of macromolecules indeed “actively maintain its structural and dynamical integrity”? This question can be answered by perturbing the system with “lesions”: selectively turning off the ability of certain macromolecules to affect the functional states of other macromolecules for the active states (Figure 5b), the sensory states (Figure 5c), and the internal states (Figure 5c). Note that all of the electrochemical effects on motion were left intact; only the subtle interaction between electrochemical states was silenced. In all three cases, such a relatively mild perturbation caused the synthetic bubble to burst instantly. This empirical result substantiates the prediction that macromolecules will affect their neighbors in order to maintain the structural integrity of the entire clump.

In this section, we have seen the emergence of a Markov blanket and resulting active inference in a random dynamical system. Functionally speaking, the simulated clump of macromolecules is probably most reminiscent of the various protein components that allow viruses to maintain their structure. We can see them as a metaphor for more extensive forms of biological self-organization. Intriguingly, Friston did not require a very “special” setup to arrive at this result in a bottom-up fashion; very little was required, in fact. This motivates our proposal to consider the recursive self-organization of Markov blankets into Markov blankets at higher levels. Each of these blankets and their internal states again constitute a unit of free-energy minimization ([Ramstead et al., 2017](#); [Sengupta et al., 2016](#)). With this in mind, we will now proceed by taking free-energy minimization “for granted” and focus instead on how this process shapes function-specificity for a single neuron (dendritic self-organization), and form-specificity at multi-cellular levels (morphogenesis).

## Simulated lesions

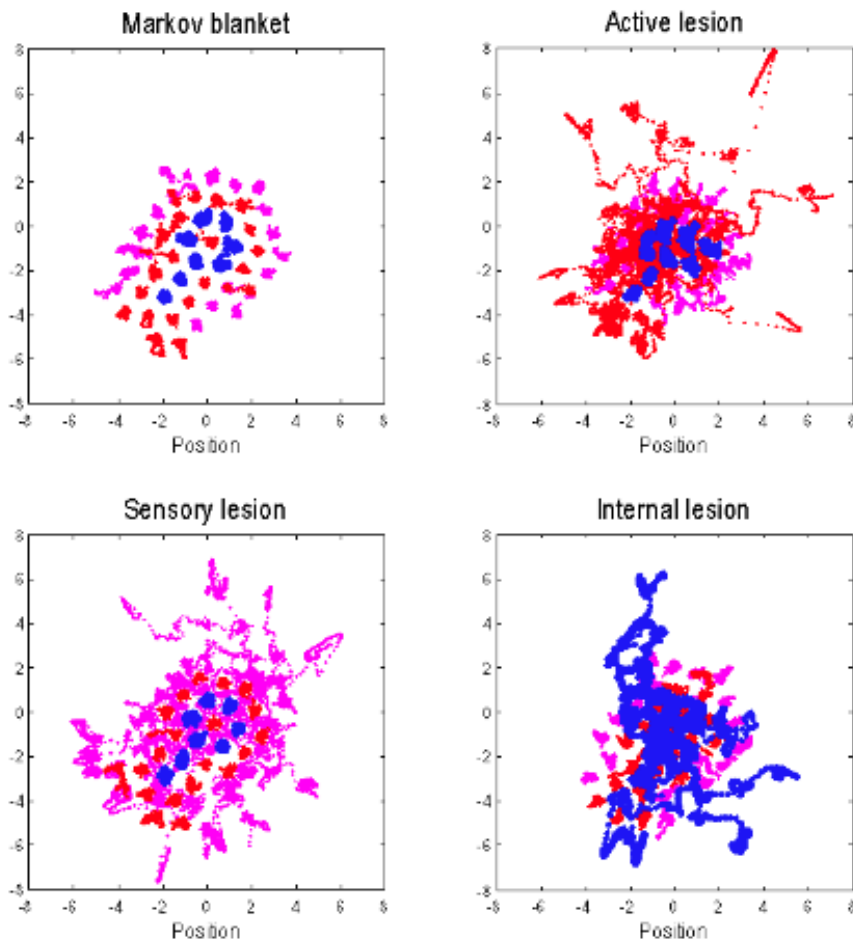


Figure 5. This figure demonstrates the self-maintaining dynamics (i.e., autopoiesis) of the clump of macromolecules, by slightly impairing the components of the emergent Markov blanket. Impaired macromolecules are rendered unable to influence the electrochemical states of other macromolecules (but all other interactions are left intact). In the top left panel, the configuration without a lesion is shown, with the internal (blue), active (red) and sensory (pink) macromolecules forming a stable configuration. In the top right panel, active macromolecules are impaired, causing them to be expelled into the exterior. In the bottom left panel, the sensory macromolecules are impaired, causing them to drift off into the exterior. In the bottom right panel, internal macromolecules are impaired, causing the entire configuration to collapse – as the internal states migrate rapidly across the Markov blanket. This figure is adapted from Friston (2013; Figure 4).

## 3.2 Dendritic self-organization

Different types of neurons code for different types of synaptic input sequences, as evidenced by their different morphologies and connections (Torben-Nielsen & Stiefel, 2009). Pyramidal neurons have been shown to engage in sequence-specific processing (Branco, Clark, & Häusser, 2010). Apparently, dendritic branches allow the dynamics within a single neuron to distinguish various sequences of input from each other. In the following, we discuss how the FEP has been used to study the emergence of such function-specificity by (Kiebel & Friston, 2011).

As stated in the introduction, under the FEP, the variational free energy represents the difference between what a biological (sub)system senses and what it expects to sense. These expectations are derived from an *implicit (generative) model* of those sensory inputs. The biological system itself *is* this model, which specifies the type of inputs it is looking for (note, once again, the inherent circularity). The minimization of free energy has been used to simulate systems that decode their sensory states and actively select the types of input they expect to sense (Kiebel, Daunizeau, & Friston, 2008). The implicit nature of these expectations and models is worth emphasizing, because it means that these Bayesian concepts do not require the system itself to be “conscious” of inferences in any way, or that these inferences need to be “explicit” and couched in propositional or linguistic terms.

A single neuron or one of its dendrites can also be understood as a biological system that engages in free-energy minimization. As we will see, this view can explain the emergence of the sequence-specific functionality of neurons towards presynaptic inputs with a certain temporal pattern. Selection of synapses occurs via synaptic gain control – synapses with low gain are pruned, and synapses with high gain stimulate the formation of synaptic connections (Lendvai et al., 2000). The concept of synaptic gain control can itself be derived from the FEP; and it can be used to capture the behavior of neuronal dynamics across multiple timescales, from fast electrochemical potentials, to variations in synaptic gain, through to slowly changing synaptic connections. In a series of simulations, Kiebel and Friston (2011) incorporated these three temporal scales in a computational model by using three levels of simultaneous free-energy minimization: a single quantity is minimized at the three scales that enclose the scale in which synaptic gain is determined. Figure 6 illustrates the type of sequence selectivity that emerged in these simulations.

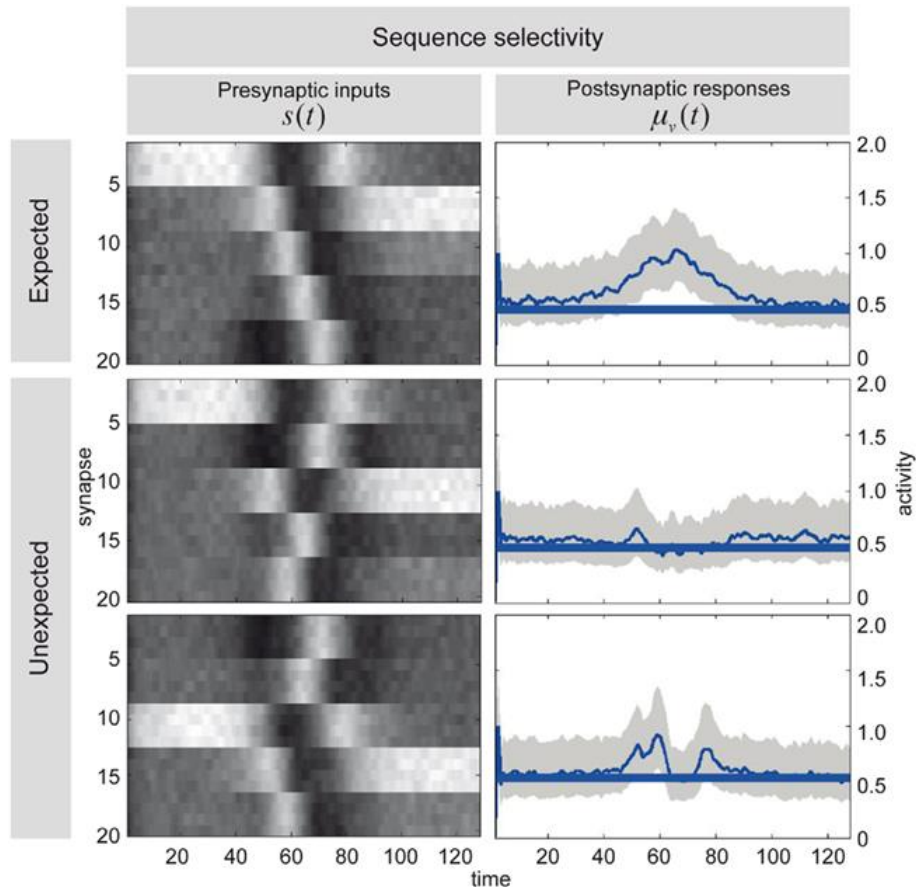


Figure 6. In this figure, we show the responses of the dendrite (right column) to three different sequences of presynaptic input (left column). The top row shows the expected sequence to which the dendrite is accustomed, showing a peak in the postsynaptic response (top right panel). The middle and bottom rows show how the dendrite responds to sequences that deviate from its expectation, with attenuated postsynaptic responses in both cases (middle and bottom right). The graded response in the bottom right panel is consistent with graded observed in neural responses to suboptimal input. Figure was taken from Kiebel and Friston (2011; Figure 7).

On the fast level of electrochemical currents, Kiebel and Friston (2011) were able to show that this free-energy-minimizing dendrite model produced emergent dynamics that were entirely consistent with data-driven models of dendritic dynamics (Gulledge, Kampa, & Stuart, 2005). Their findings showed that such active dendritic dynamics are a self-organizing function of this particular biological system under the FEP. The slow dynamics of the dendrite - rearranging the synaptic connections over time - is incorporated in the model as a form of Bayesian model selection. The various connections are essentially producing evidence for their own efficacy with varying degrees of success, instantiating a process of selection over time. Selection occurs stochastically, allowing for completely non-efficacious configurations, but also rendering the routine better equipped to escape local (sub-optimal) minima. In Section 5, we discuss how a similar kind of dynamics governs evolution by natural selection. Notably, a similar type of model selection is also believed to drive the fine-tuning of entire neural networks, which has been broadly conceptualized as *neural Darwinism* (Edelman, 1987).



In this sub-cellular example, the dendrite is minimizing free energy to improve: (i) its beliefs about presynaptic input sequences on short timescales; (ii) its beliefs about synaptic gain (or precision); and (iii) its implicit model of the input sequences over longer timescales. In this way, the dendrite adjusts its prior beliefs about the type of sequences it expects to observe, which results in the observed selective sensitivity. The sampling method of the dendrite is being adjusted over time, which boils down to a type of active inference. In the following, we consider how a group of cells (free-energy-minimizing units) can self-organize into larger structures; namely, organs.

### 3.3 Morphogenesis

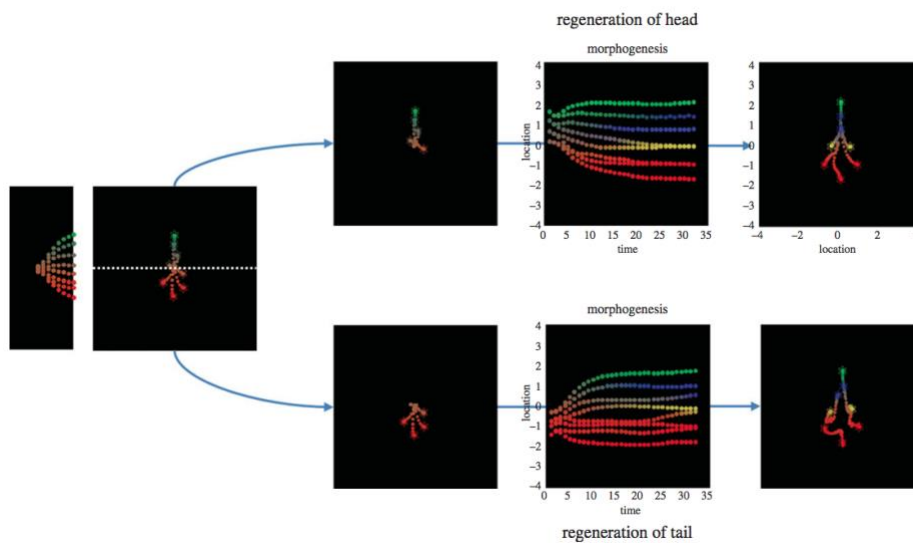
Now that we have established cells as units of free-energy minimization, we can consider how adaptive self-organization occurs under collective active inference; i.e., the group dynamics of cellular ensembles (Friston et al., 2015). An important example is the emergence and maintenance of the large-scale shape and function of entire subsystems (e.g., organs). How can cells at microscales coordinate to form pre-defined large-scale structures; e.g., during embryonic development? Or, at later stages, how can creatures like salamanders regenerate entire limbs and organs? It is an essential question for biology, both in development and throughout evolution, to consider how cellular ensembles control exact large-scale outcomes in order to allow for specific functions to emerge (e.g., brain or liver function). Insights into this issue are particularly crucial to medicine and bioengineering.

As we will see, collective active inference can explain the self-organization of an ensemble of cells to generate entire organs (i.e., morphogenesis). The most pressing difficulty here is that organs will only function if they have a highly specific, predefined form, for which unguided pattern formation is insufficient. Since any one cell only has access to the signals reaching its boundary, it would seem that it can only infer its location and differentiate once the other cells have already migrated to their respective target positions and differentiated accordingly. However, that requirement cannot be reached if those other cells themselves are unable to determine their own target positions.

This inherently circular problem of organ formation can be solved through active inference, if we assume that every (pluripotential) cell starts with a generative model of the entire ensemble. In this way, every cell can generate predictions about the sensory inputs it expects to encounter at any location in the target configuration. As with stem-cells, all cells start out in nearly identical states, with the same generative model and the ability to differentiate; that is, transition towards any role in the eventual organ. As each individual cell starts minimizing free energy, the entire ensemble will converge towards its global free-energy minimum. By virtue of their common generative model, this global minimum is approached when the ensemble closes in on the target shape and function of the organ. Each cell will gradually infer its own place and behave accordingly, while, crucially, helping other cells to infer their place in the process. Such self-assembly will also serve to maintain the configuration and, in the case of damage to the organ, restore it.

In order to substantiate this account, Friston and colleagues (2015) conducted simulations of cell migration and differentiation in a relatively minimalistic sense. Each cell possessed a generative model, the parameters of which were determined “genetically” (they were inherited or pre-specified), which prescribed to each cell how to act (i.e., what signals to emit) given a particular place within the organ. Hence, cells exchanged signals with each other in order to infer their respective place and role in the ensemble. The upshot of this is that every cell has a probabilistic grasp on its location and emits signals accordingly, providing information for the other cells to improve their own inferences. That relation between the beliefs of a cell concerning its place and the signals it transmits to other cells could be an elegant metaphor for epigenetic processes. Figure 7 serves to illustrate the simulation results for a configuration of a relatively small number of cells. It shows both the differentiation process and the reorganization of the ensemble after two different large lesions.

We have illustrated how ensembles of free-energy-minimizing units (cells) that operate with the same generative model can self-organize into pre-determined structures (organs). This allows us to understand how an intricate functional structure like the brain can be produced by the (epi)genetic information transmitted at conception. This treatment has prepared us for a discussion of the brain, entire organisms, and their interactions. Interestingly, we will see a similar sort of dynamics emerge in the interaction between multiple organisms: a shared generative model allows for the emergence of communication and cultural dynamics.



**Figure 7.** This figure shows both the differentiation of eight stem cells to form an “organ” (on the left) as well as the regenerative response of the configuration to two large lesions (on the right). In the top three panels on the right, the “head” (consisting of red cells) is severed and the remaining cells are doubled to maintain the same number. On the bottom, the same operation is performed on the “tail” (consisting of green cells). Both show that the pattern is successfully recovered. Figure taken from Friston et al. (2015; Figure 4).

## 4. Mesoscale: Organisms and Their Interactions

### 4.1. The brain

At this point, we arrive at the level of organization involving animals and the interactions between them. We would be remiss if we would not reserve a few words for the animal brain in particular. Its organization and functional dynamics could be understood in terms of the examples of free-energy minimization treated thus far. The brain exhibits a layered and modular structure, instantiated through morphogenesis (Section 3.3). We suggest this organization of the brain has been selected for throughout evolution (Section 5.1) because it enables the assembly and maintenance of hierarchical generative models (Badcock, Friston, & Ramstead; under review; Friston, 2010). In our environments, there is an abundance of hierarchical inference problems. For example, in the case of natural images, the integration of large numbers of features is required in order to identify objects under countless possible lighting conditions and rotations in space. In computational neuroscience, free-energy minimization has led to the development of models engaging in hierarchical predictive processing that successfully capture the functioning of the brain (Adams, Bauer, Pinotsis, & Friston, 2016). The brain is thus viewed as an active inference machine (Clark, 2015), specialized for complex inferences requiring hierarchical generative models. It would not be an overstatement to say that it is the most complex adaptive system known to mankind, as it continuously bridges the scale space from genes to single dendrites up to organismic and societal levels (Ramstead, Badcock, & Friston, 2018). Under the FEP, the brain essentially functions like its many lower levels of organization: it predicts sensory states from its internal model(s) of how those sensory states are caused (see the hallmark paper on the FEP by Friston, 2010). It minimizes the discrepancies between its expectations and actual sensory states by modifying its implicit beliefs (i.e., perception) or by acting on its environment (i.e., behavior). The inferential power of the hierarchical organization of the brain can be well illustrated by studying how it generates predictions about another hierarchical dynamical system: namely, another organism. We choose not to focus on how the brain instantiates bare forms of perception and action, but on how two bird brains are coupled through birdsong. This will serve as an informative example of hierarchical inferential dynamics enabled by free-energy minimization.

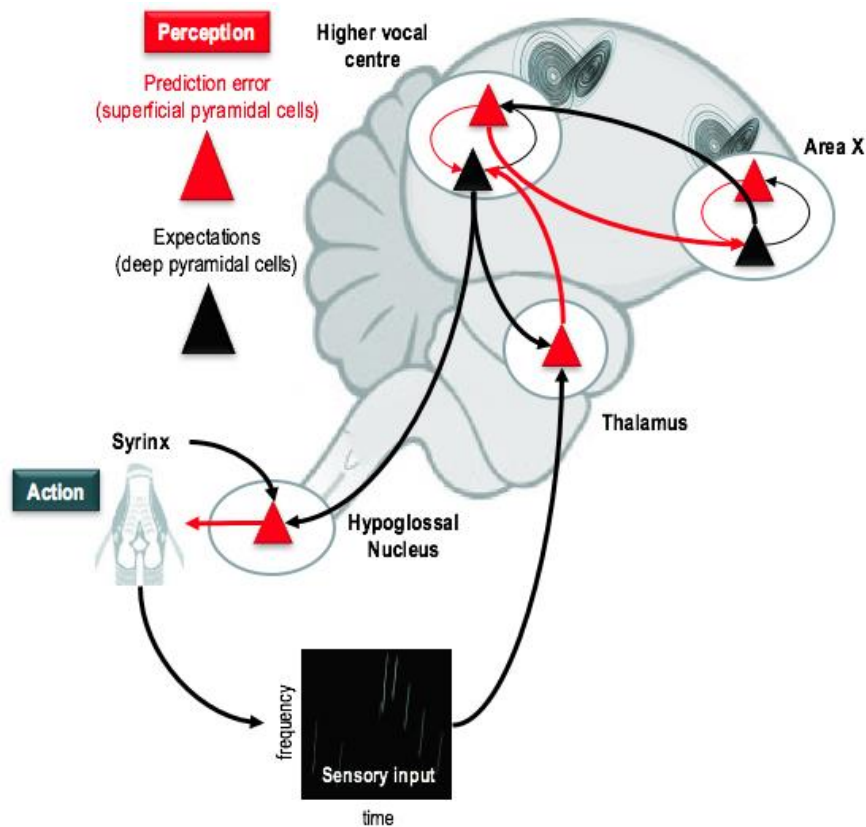
### 4.2 Birdsong as a model of dialogue

When two dynamical systems are coupled to each other, a form of synchronization usually occurs. This was first reported by (Huygens, 1673), who studied the synchronization of pendulums hanging from a beam, through which they influenced each other very slightly. Because both pendulums operate in the same way, even the minimal information transmitted by the beam is enough to completely synchronize them. In a similar way, coupled brains can, by virtue of their similar internal dynamics, achieve generalized synchrony. Such synchrony allows for these systems to predict one another with very high precision. In the case of identical internal models, identical synchronization is achieved (similar to the case of the pendulums). The more dissimilar the internal models of two organisms, the less synchronization will occur between their internal states and, consequently, the less accurate their predictions will be about each other's actions. Without environmental constraints, coupled organisms will tend to move towards the free-energy

minimum of identical synchronization. In other words, they end up forming a model of each other. Through such coupling, organisms can “program” each other towards a common internal model; namely, they end up speaking the same “language” (in an abstract sense).

The way in which dynamical coupling gives rise to generalized synchrony in pendulums can thus be applied to the fine-tuning of hierarchical internal models that generate predictions. Such learning was addressed by Frith and Friston (2015), which is the focus of this section. The authors demonstrated how organisms can come to interpret each other’s actions simply by adjusting their internal models to minimize free energy. Importantly, free energy can be evaluated without these organisms ever knowing exactly what is happening beneath the Markov blanket of the other. It relates to the central problem of hermeneutics: how do we infer the intention behind an utterance, when we only have access to the utterance itself? In the following, we discuss simulations by Frith and Friston that are an abstract representation or metaphor of communication between organisms, based on the mathematical machinery of complex dynamic systems. Synthetic birdsong is used for this demonstration, but it is not meant to represent actual linguistic processes. The authors merely aimed to study dynamic coupling via complex action patterns, which are themselves without meaning or syntax. However, it is worth noting that other researchers have applied hierarchical-predictive processing to language (e.g., Hickok, 2013) and auditory processing (Arnal, Wyart, & Giraud, 2011).

In order to simulate birdsong-like behavior, Friston and Frith (2015) constructed a hierarchical processing architecture, which is shown in Figure 8 (overlaid on analogous neuroanatomical structures of the bird brain). Free-energy minimization is achieved through recurrent connections between different levels of the hierarchy, each of which possesses its own generative model. Each level generates its own expectations about how sensory inputs are caused, which are passed downward as predictions. Each level (except the highest one) therefore receives (top-down) predictions to compare with its own expectations. The difference is the prediction error, which is passed back to the higher levels in a bottom-up fashion in order to improve future predictions. Experimental findings appear to support such an architecture. For example, it has been suggested that superficial pyramidal cells are involved in calculating prediction errors and passing them upwards and that deep pyramidal cells pass the expectations of each level to the one below in the form of predictions (Bastos et al., 2012). In this hierarchy, the predictions of the lowest level are essentially those generating motor commands and corollary discharge.



**Figure 8.** This figure illustrates schematically the hierarchical predictive-processing architecture of the songbirds, overlaid on (possibly) analogous neuroanatomical structures of an actual bird brain. Red arrows indicate the flow of information about the prediction errors, transmitted by superficial pyramidal cells (red triangles). Black arrows indicate the flow of information about the expectations on each level, transmitted by deep pyramidal cells (black triangles). Area X transmits predictions to the higher vocal center, which generates drives the hypoglossal nucleus to generate a vocal response (via the syrinx) as well as the thalamus to generate the corollary discharge. Adapted from Friston and Frith (2015; Figure 1).

The simulations of Friston and Frith (2015) showed that two of these synthetic bird brains became coupled through their vocalizations during a turn-taking exercise, providing clear evidence of generalized synchrony. These dynamics occurred at the free-energy minimum of the two coupled systems. Importantly, a high degree of synchronization was achieved because both systems started out with a similar architecture (or neuroanatomy), by virtue of being birds. Both of these birds were simply predicting their own sensory states, using a hierarchical composition of hidden states. The final product emerged in their “dialogue”, so both of their hidden states had essentially come to represent their shared expectations. This meant the only thing the birds had to infer was which one of them was singing (i.e., agency). This inference enabled them to either attend or ignore the sensory consequences of their action; depending upon whether they were listening or singing.

Perhaps a more interesting – and realistic – case is when two birds are in some ways dissimilar to each other, rendering their dyad asymmetric. One of the birds was given a mild handicap by reducing its responsiveness to top-down predictions, which also hampered the quality of its vocalizations. As shown in Figure 9, this adjustment allowed a type of scaffolding dynamics to emerge, in which the more proficient bird simplified its own

vocalizations in order to accommodate the shortcomings of the other bird. Through this process, they reached nearly identical synchronization, solving the hermeneutical problem in the process (so to speak). Interestingly, this kind of demonstration is analogous to scaffolding techniques used in teaching, in which the teacher optimizes learning by lowering his or her level of instruction close to, but slightly above, that of the student. For reference, Figure 9 also includes a simulation in which the birds are disconnected from each other, showing how heavily the richness of their vocalizations depended on the presence of another bird. When the birds were alone, they started learning from the silence around them to become silent themselves. It shows that, in some way, the teacher was actually learning from the student too.

Although this coupled setup was rather ad hoc, it can be seen as a step towards understanding the development and evolution of social life. Through generalized synchrony, one could efficiently infer the sensations and action goals of others, a crucial aspect of higher cognitive functions. Important examples are vicarious learning (learning by watching others), empathy (inferring others' feelings), and theory of mind (inferring others' inferences). A form of generalized synchrony appears to underlie mirror neuron activity in animal brains – mirror neurons not only fire during certain actions or sensations, but also when observing a conspecific performing or experiencing similar actions or sensations (Friston, Mattout, & Kilner, 2011; Kilner, Friston, & Frith, 2007). This type of associative mirroring of neural responses appears to be similar to the generalized synchrony exemplified in the above birdsong simulations. It has been argued that mirror neurons are an associative byproduct of action-understanding and empathy (Hickock, 2009; Cook et al., 2017). In future work, studies that investigate the ways in which free-energy minimization leads to generalized synchrony between organisms might help explain observations of mirror neuron activity.

In this section, we have discussed the sort of learning dynamics that emerge when two hierarchically structured, free-energy-minimizing (bird-like) organisms interact. Once again, circular relationships are involved, now in the context of communication and generalized synchrony, resulting in the emergence of shared expectations. In the following, we discuss how shared expectations and narratives shape human cultural dynamics.

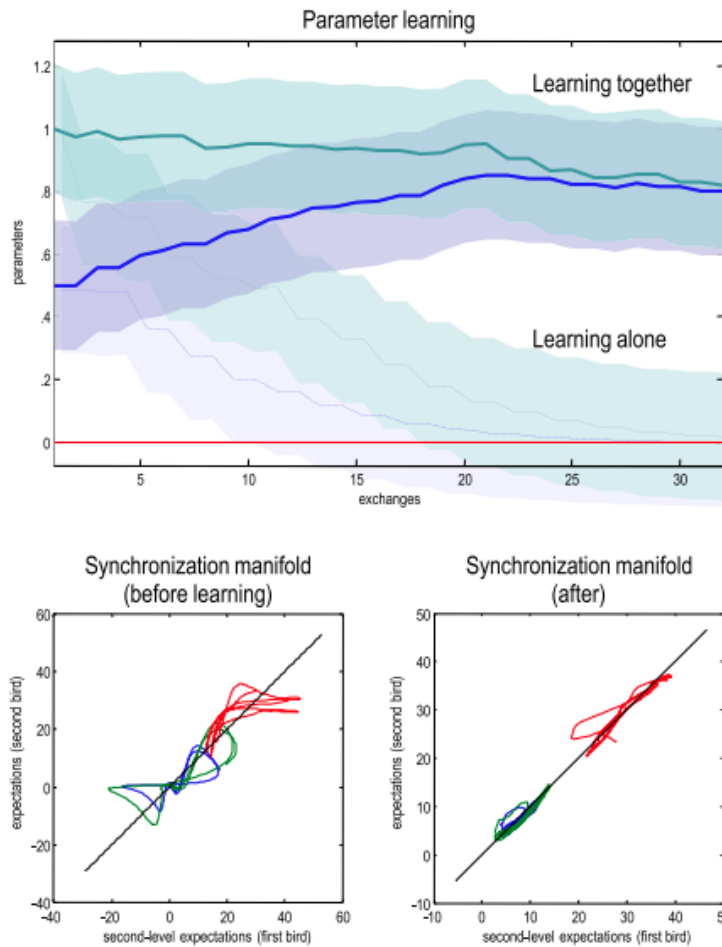


Figure 9. This figure illustrates both the learning of two coupled birds (top panel) and the generalized synchrony reached after their exchanges (bottom two panels). The top panel shows the changes in both birds' (posterior) beliefs about a parameter that controls the prosody (or richness) of their vocalizations over a number of exchanges (birds taking turns; either singing or listening). The proficient bird is shown in green, the less proficient one in blue. 90%-confidence intervals over this parameter are indicated by the shaded areas. The bottom panel shows the degree of synchronization between the expectations of the birds about three hierarchical, dynamic states that drive the singing behavior (red, green, blue), both before (left) and after (right) their exchanges. Since the x-axis shows the expectations of the less proficient (first) bird and the y-axis those of the more proficient (second) bird, synchronization is achieved on the line  $x=y$ . Figure taken from Friston and Frith (2015; Figure 8).



## 4.3 Cultural ensembles

So far, we have seen – across various scales – how biological systems come to embody an implicit model of their environment through active inference. The emphasis on organism-environment coupling is inherent to the free-energy principle, which plays very well into another framework that has recently gained traction among researchers, *ecological and embodied approaches to cognition* (Bruineberg & Rietveld, 2014; Chemero, 2009; Gibson, 1979; Kirchhoff, 2015; 2017). In this section, we discuss an example of recent efforts to connect these frameworks in the context of human cultural dynamics by (Maxwell J. D. Ramstead, Veissière, & Kirmayer, 2016)

A synthesis between the free-energy principle and the ecological approach allows these approaches to benefit from each other's insights and research. From the conceptual toolbox of ecological cognition, we introduce the notion of *affordances*. Affordances are possibilities for engagement through action and perception that are enabled by the relationship between the environment and the abilities of the organism in question. Under the FEP, an organism acts on its environment in order to bring about its preferred (expected) sensory outcomes (Bruineberg, 2018). In this way, free-energy minimization specifies the most likely trajectories of organisms in their landscape of affordances.

Ramstead, Veissière, and Kirmayer (2016) made the distinction between natural and cultural affordances. Affordances of the first kind are derived directly from the environment (e.g., walking) and only require minimal social learning; while those of the second kind are derived from the shared expectations inherent to the (sub)culture in question (e.g., language) and require more extensive social scaffolding to be acquired and used effectively. The previous section illustrated how shared expectations can emerge from interactions between two organisms. In the case of culture, we generalize this notion to a population of interacting organisms united by one common set of shared expectations, which in turn shape the various possibilities for interaction: namely, cultural affordances. Of course, the distinction is not absolute; natural and conventional affordances are more like the opposite ends of a spectrum of affordances. For example, in many cases, conventional rules simply act to constrain natural affordances (e.g., driving on the wrong side of the road).

Researchers developing the concept of affordances emphasize that agents use sensory information for affordances, without requiring explicit representations of the affordances themselves (van Dijk et al., 2015). This minimalistic view sits well with active inference, given that statistical terms are seen to be implicit (as we noted earlier). Expressed otherwise, internal models are implicitly instantiated by the dynamics themselves. For example, in the hierarchical architecture introduced in Section 4.1-2, free-energy minimization occurs locally on each level of the hierarchy, based only on the neural signals incoming from adjacent levels. None of these levels necessarily requires “meta-cognitive” contextual information about the hierarchical internal model.

So how do humans become so proficient in leveraging this field of (implicit) cultural affordances? Under a hierarchical predictive-processing architecture, any level can modulate expectations at the level immediately below it, thereby modulating which types of input that lower layer is sensitive to. Such prior expectations can implement a gating mechanism, which has been proposed to explain attention. In principle, cultural affordances

could then be learned by fine-tuning these priors to induce selective attention, which constrains the field of all possible affordances. Effectively, this can be arrived at by extending the modelling strategy for morphogenesis of Section 3.3, by equipping all cultural agents with the same cultural priors. Such culture-specific fine-tuning of internal models can occur through the type of generalized synchrony discussed in Section 4.2. Shared expectations that emerge from collective free-energy minimization induce “regimes of shared attention” that guide and constrain social practices, which in turn shape those expectations (Ramstead et al., 2016).

Under this view, social norms can be cast as shared “solutions” arrived at, and learned through, the collective free-energy minimization of people within a particular culture (Colombo, 2014). The shared aspect of social norms reflects a certain degree of synchronization between people within a given (sub)culture, allowing them to produce more accurate inferences about each other’s internal states. For example, it is much easier to predict the actions of, and empathize with, somebody from your own (sub)culture than somebody from an alien one. This emergent view of social norms and practices corresponds well with *social constructivism*, a well-established framework in sociology which emphasizes that human development is socially embedded and human narratives are constructed through interaction with others (Berger & Luckmann, 1966). We suggest that the free-energy principle can undergird social constructivism by explaining how shared cultural narratives can emerge from, and are learned through, human interactions.

Finally, the shared aspect of cultural affordances suggests that human social capacities emerged not “just” because of more advanced hierarchical internal models, because one’s grip on cultural affordances is learned only through interaction with other humans within that culture. The converse would therefore seem more likely: our processing hierarchy has been optimized through evolution in order to keep up with the growing demands of the early social practices of our primate ancestors. This is a prime example of evolution through both natural selection *and* niche construction, which we discuss in the following section.

# 5. Macroscale: Species as Families of Model-Niche Pairings

## 5.1 Evolution as Bayesian model selection

We are now prepared to address one of the three central topics of this volume - the evolution of species, which provides the context in which the smaller temporal scales of adaptive self-organization are embedded. We assume familiarity with evolutionary theory, so we do not completely hash out the basic concepts of evolution, but rather explore how these concepts can be understood as free-energy minimization at the species level. In particular, we discuss evolution as a form of Bayesian model selection.

In our treatment thus far, we have assumed biological systems to be ergodic. Ergodicity implies that a system only resides in a limited number of states over time, which makes probabilistic inferences (and hence active inference) possible. Of course, real biological systems are only locally ergodic. Throughout the development of an organism, various states are pruned away and new ones are unlocked, sometimes quite radically (e.g., a caterpillar becoming a butterfly). Eventually, death involves a divergence of possible configurations - a complete breakdown of ergodicity (from the perspective of the phenotype). The complex adaptive systems, we refer to as organisms, do not maintain their structure and function forever: indeed, in a changing environment, the emergent Markov blanket of Section 3.1 would eventually be destroyed. In the beginning of evolutionary history, this (perhaps inevitable) disintegration has been overcome through the emergence of the ability to reproduce. Reproduction is an adaptive capacity that allows genetic, epigenetic, and non-genetic information to be transmitted to descendants along with small variations, constraining the self-organizing dynamics that specify the form and function of their internal models for active inference. Through inheritance and the subsequent experiences of organisms, every new generation introduces variations of the internal models of their parent population. Inherited aspects of these internal models can be realized in various ways, which we discuss now.

In Section 3.3, we saw how the large-scale shape and function of organs can be fine-tuned through the initial internal models of stem cells. Such processes can bring about the hierarchical organization of the animal brain, which in turn allows for hierarchical internal models, as discussed in Section 4.1-2. Besides the overall hierarchical structure of internal models, another type of heritable modulating mechanism could be instantiated through adaptive priors (within a given brain organization) that predispose the organism to learning certain types of structures ([Friston, Thornton, & Clark, 2012](#); [Ramstead et al., 2017](#)). For example, humans appear to have an innate disposition for the acquisition of language. Another very important form of evolutionary preparedness is the inborn affective value of various types of stimuli. In the context of free-energy minimization, an innate tendency to approach or avoid certain situations could be implemented implicitly through prior preferences over sensory inputs. Internal models can be adapted to tweak the expected free energy under various sensations, without (strong) reliance on learning through experience. For example, we all respond with disgust to the smell of rotten eggs without ever having experienced hydrogen sulfide poisoning. On a more positive note, we all tend to enjoy the

taste of sweet and fat-rich food (a tendency skillfully exploited by modern fast-food chains). There are also examples of complex stimuli that are known to have an innate affective value. For example, all mammals appear to be predisposed towards developing a fear of snakes (Badcock, Ploeger, & Allen, 2016). Captive-born lemurs and macaques learn to fear snakes faster than other types of equally rich stimuli (Weiss, Brandl, & Frynta, 2015). This finding has led some to suggest that snake-like reptiles used to be a large threat to the survival of mammals in an early stage of evolutionary history. Under the free-energy principle, innate preferences over inputs are not limited to the lowest (sensory) level of the predictive-processing hierarchy. Preferences over inputs can also apply to the incoming (sensory-driven) signals on higher levels, which could explain the innate affective value of highly complex stimuli like snakes. Again, on a more positive note, the same mechanism can also explain the positive experience of “cuteness” invoked by the bodily proportions of babies (and, probably an evolutionarily “accidental” corollary, puppies and kittens). Indeed, the important role of adaptive priors in active inference has even been leveraged to explain highly complex human phenomena, such as our capacity for depression (Badcock et al, 2017).

**Now that we have specified the ways in which evolutionary preparedness can be realized through internal models, we can consider the selection process itself. Natural selection is underwritten by differentials in adaptive fitness. Whatever traits are most suited to ensuring the survival and procreation of individual are most likely to be transmitted (genetically and epigenetically). Consequently, these traits will occur more frequently in subsequent generations. Constrained by the transmission of (epi)genetic information to the next generation, natural selection acts primarily on individuals (i.e., individual fitness); although it can also occur through an individual’s contribution to the survival and reproductive success of others, especially close relatives (i.e., kin selection and inclusive fitness; Dawkins, 1976; Hamilton, 1964; Maynard Smith, 1964; Orgel & Crick, 1980). Notably, the evolutionary success of a species depends strongly on the amount of (epi)genetic variation present in the populations that constitute the species (evolutionary resilience; e.g., Sgrò, Lowe, & Hoffmann, 2011). Such variation increases the likelihood of the presence of individuals with high fitness under new, challenging circumstances. Every individual represents an attempt to transmit its (epi)genetic makeup, such that natural selection effectively produces a stochastic gradient ascent on the expected fitness of the population (as employed in machine learning by, e.g., Yi et al., 2009).**

The FEP provides a framework to predict adaptive fitness from first principles, while also taking into account organism-environment interactions. In effect, maximizing the adaptive fitness of a population is likely achieved by minimizing its collective free energy, which tracks the goodness of fit (or complementarity) between the states of a species, and the states of its niche. Accordingly, individuals that are well suited for survival are those that minimize free energy efficiently. Generation by generation, the adaptivity of individual organisms can be evaluated by the negative time-average of free energy (i.e., a lower bound on entropy). Since free energy itself is evaluated using the internal model of a specimen, a comparison only has predictive value for adaptation if the family of internal models and the niche under consideration are similar. The strictly local utility of this comparison illustrates

the incremental nature of evolution.<sup>2</sup> For example, using free energy as a metric, one could score the complementarity between, say, a bacterium and its niche, and a human being and its niche. In short, free energy could provide a universal proxy for adaptive fitness that could be applied to both viruses and vegans. At the same time, free-energy minimization is achieved through gradient descent, which means in this context that it is quintessentially species – or model – specific.

As noted in the introduction, minimizing free energy is formally equivalent to maximizing Bayesian model evidence; that is, the likelihood of the internal model being true or apt, given the organism's environment. Therefore, we are now in a position to interpret processes of adaptation as collecting Bayesian model evidence and, by extension, to cast natural selection as a form of Bayesian model selection (see also Campbell, 2016). On this view, creatures are naturally selected according to how well their internal generative models fit with the environment. Of course, this picture becomes more complicated in the case of organisms that interact with each other to increase total fitness (i.e., decrease collective free energy). These multiple organisms are not only "fitting" their shared environment but also each other, generating shared expectations in the process (as seen in Section 4.2). By virtue of the inherited directives for their internal models (i.e., adaptive priors), which have been shaped by natural selection, organisms minimize their free energy locally over their own (relatively short) lives in ways that *also* help their descendants (e.g., parents nurturing their children) and close relatives (i.e., kin selection). Local (organismic) free-energy-minimizing dynamics are structured in such a way that they collectively move towards a (population-level) free-energy minimum. This type of relationship between local and global dynamics is analogous to the predictive-processing hierarchy in the brain, as described in Sections 4.1-2. Every layer in the hierarchy minimizes its own free energy (locally), in such a way that it also helps the hierarchy as a whole move towards its free-energy minimum (globally).

Thus far, our discussion of evolution has yet to explore how organisms shape their own environment, which can also become part of the inheritance they leave behind for their descendants. Such niche construction – and implicit legacy – is the focus of the final section.

## 5.2 Niche construction

Niche construction is the process by which organisms modify their environment through their normal bioregulatory activity (Odling-Smee, Laland, & Feldman 2003). It encompasses *all* modifications, from the induction of a layer of moist air around homeothermic organisms, to the construction of complex environments like cities by human beings. Like perception and action, niche construction is ubiquitous in living systems. Indeed, it is a direct corollary of active inference – organisms attune the statistical structure of their environment to their probabilistic expectations by acting in a way that is guided by those expectations. We have seen that perception enables the organism to infer sensory causes; action places an upper bound on surprise by generating expected changes in the sensorium. The variational free-

---

<sup>2</sup> From a technical point of view, the extensive nature of free energy means that the sum of the free energy of the parts is equal to the free energy of the sum; so what is 'good' locally is good globally. This extensive characteristic implies that minimizing free energy over time is analogous to the Hamiltonian Principle of Least Action – because Action is the integral of energy over time.

energy approach to niche construction exploits the symmetry in the Markov blanket formalism; namely, between internal and external states mediated by the blanket states (i.e., the fact that action engenders modifications of the local environment, which embeds sensory causes). In this section, we explore the role of such ecological modifications with regard to evolution.

By virtue of ergodicity, an organism may be defined as the most *likely* set of physiological and behavioral states for any given set of environmental states. The coupling between these states then constitutes the entire organism-environment state space. As stated in Section 5.1, adaptivity is a feature of an organism-environment system, not just of organisms themselves (e.g., gills are adaptive for water-bound organisms; lungs for those dwelling on land). Negative variational free energy can be seen as a measure of adaptivity (as in Section 5.1); either for individual organisms and their niche, or for larger ensembles like groups and species and their environment. It tracks the extent to which the statistical organization of an organism's physiological and behavioral states transcribe the statistical organization of the states of its environment.

Among those states, some pertain to the internal organization of the organism. These are fast, fluctuating states, like synaptic connections and neuromodulatory gating patterns. Some other states pertain to the external, visible organization of the organism (i.e., phenotypic states). These are more slowly fluctuating quantities, like behavioral patterns and morphological features (i.e., phenotypic traits). States of the environment themselves can be interpreted as part of those slowly fluctuating states. The level of adaptivity among slowly and rapidly fluctuating states depends on the interplay between variational optimization processes spanning different spatiotemporal scales, ranging from natural selection (Bayesian model selection), through to development and learning (active inference).

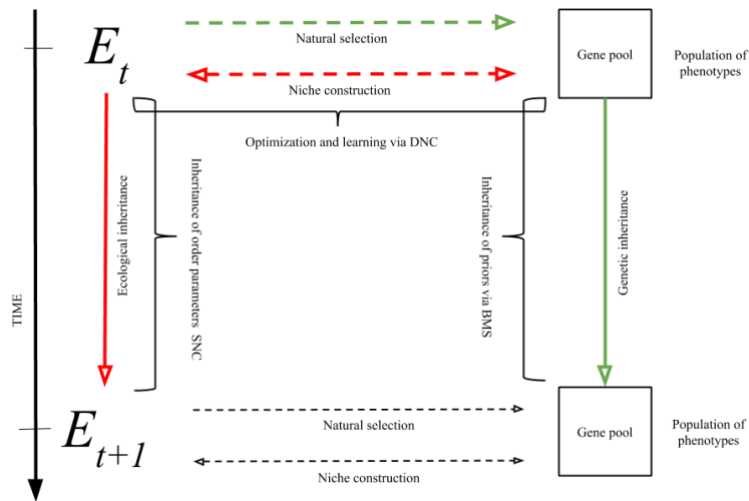
At this point, the notion of action and environmental modifications become important. Because action fulfills sensory expectations (e.g., adaptive priors concerning viable states, like body temperature), it can change, implicitly, the statistics of the niche so as to make them consistent with the sensory expectations of an organism. In other words, niche construction fits sensory causes to sensory expectations, and reciprocally, fits sensory expectations to sensory causes. Under active inference, niche construction is crucial in allowing for optimization across the scales of the spatiotemporal hierarchy: the more slowly changing parameters embodied by or encoded in the physical features are optimized through niche construction, and in return act as a kind of developmental driver by channeling adaptive behavior and phenotypic accommodation (Bruineberg, 2018; [Constant et al., 2018](#)). In a sense, the robustness of living systems is inherited from the regularity and stability of their more slowly changing eco-niche.

Adaptation is often conceived of as a one way process, by which natural selection shapes organisms under the pressures of their environmental conditions: a view sometimes called 'externalism' in the context of natural selection ([Godfrey-Smith 1996](#)). As we considered in Section 5.1, those pressures pose challenges, the resolution of which rests on the retention of organisms that are best suited to gain differential fitness. Under the FEP, this corresponds to the selection of (constraints to) internal models that are most suited to minimize free energy.

The niche construction perspective involves a complementary view of adaptation, in which internal factors, like the states of organisms, also play an evolutionarily significant role in their adaptation. Organisms generate feedback interactions with their environment, which can steer their own evolutionary trajectories, not to mention those of other species (J. Odling-Smee et al., 2003). These can generate new challenges, requiring the deployment of novel traits and behaviors in order to resolve them.

Recursive processes in niche construction impact two different, yet overlapping spatiotemporal scales: development and natural selection (Stotz 2017). At the level of development, niche construction modifies the environmental inputs to an organism's development, along with those to its offspring (e.g., through parental care). Such modifications often involve making the environment congruent with the expectations of the organism (Constant et al., 2018). At the level of evolution, niche construction functions as a strategy to modify the selection pressures afforded by the environment, thereby impacting the adaptive fitness of future generations. For instance, as a natural consequence of dam building, beaver kits inherit ecological resources like dam remains that, in turn, support the typical life cycle of beavers (Naiman, Johnston, & Kelley 1988).

We can thus see how niche construction leads to the inheritance of environmentally transmitted information (as opposed to information transmitted through reproduction) that, throughout ontogeny, helps the organism minimize its uncertainty about the states of its environment that are likely to provide a fitness advantage (e.g., palm nut residues that guide the learning of food exploitation techniques in capuchin monkeys; Fragaszy, 2011; Fragaszy et al., 2017). Such information is known as algorithmic information, which is an important source of non-genetic inheritance (Odling-Smee et al., 2013). In the context of free-energy minimization, algorithmic information enables the organism to maximize mutual information between the model it has genetically inherited (i.e., its adaptive priors), and the causal states of the environment that it has inherited ecologically. Indeed, in virtue of the symmetrical statistical dependencies across the Markov blanket of any phenotype (or ensemble of phenotypes), one can also regard the environment as entailing a generative model of the phenotypes (or ensemble) to which it plays host. Again, we see a circular causality that can be operationalised by noting that the free energy of a creature is complemented with a (conjugate) free energy of its environment; where the active states of the creature become the sensory states of the environment (and vice versa). If the free energy of either forms a proxy for adaptive fitness, we have a formal measure of "fitness" that can be applied to both phenotype and econiche. From an evolutionary perspective, this means that the environment will appear to be subject to selective pressure. In summary, the FEP therefore undergirds niche construction theory by providing: (i) a principled measure of fitness that is optimized across spatiotemporal scales; and (ii) a computational framework to reflect on ecological inheritance.



**Figure X.** Adaptation under the FEP. This schematic – inspired by Odling-Smee & Laland (2000) – illustrates the evolutionary processes covered thus far in this chapter (colored arrows). These conspire in real time  $t$  to secure the adaptation of future generation at time  $t+1$ . Ecological inheritance via selective niche construction (SNC, top-down, red full arrow) is interpreted as the transmission of environmental components that support variational updates (learning) in development (e.g., phenotypic accommodation). The FEP interprets genetic inheritance as Bayesian model selection (BMS, top-down, green full arrow), which leads to the inheritance of model components, selected on the basis of their ability to maximize adaptive value (negative surprise). Inherited priors are those predictable from the organism’s ancestors’ ability to cope with the environment, in the sense of attaining free-energy minima (or the neighboring of a limited repertoire of physiological and behavioral states). Niche construction over development (DNC, lateral, bidirectional red dotted arrow) is described in terms of model optimization via active inference, and entails ecological inheritance. Note that niche construction in development causes the symmetry between the organisms and the niche they inhabit, hence the bidirectional arrow.

## Conclusion

In this chapter, we have demonstrated how the FEP can be applied to understand adaptive, biological self-organization across spatiotemporal scales. Free-energy minimization implies active inference, which in turn allows biological systems to actively maintain their structure and function. We have discussed how Markov blankets, the basic unit of free-energy minimization and requirement for active inference, can emerge by themselves from a primordial soup. Across the manifold scales considered herein, similar processes of adaptive self-organization recurred in various ways – just as Bayesian model selection gives rise to sequence-specificity in a single dendrite (Section 3.2), it also shapes entire neural networks (neural Darwinism), and can be used to understand natural selection (Section 5.1). Shared internal models allow for the organization of many cells into entire organs (Section 3.3), but they also allow for the emergence and continuation of dialogue (Section 4.2) and culture (Section 4.3). Local optimization at separate levels of the hierarchical brain also enables system-wide free-energy minimization (Sections 4.1-2), while individual free-energy-minimizing organisms contribute to the adaptive fitness of an entire species (Section 5.1). Just as organisms can carve out expectations in each other’s internal models through interactions (Section 4.2), they can construct niches in their environment that sculpts the models of their descendants (Section 5.2). All of these interconnected examples serve to illustrate how the FEP has the potential to provide a unifying framework for the multi-scale complexity of life. Our intention is not to replace existing theoretical frameworks, but rather, to provide an underlying, quantifiable description from first principles that can be used to



integrate and coordinate such frameworks. For example, along the way, we have discussed embodied cognition, social constructivism, evolutionary theory, and niche construction. A unifying theoretical description can provide support for these various frameworks and allow them to benefit from the mathematical machinery of the FEP.

# References

- Adams, R. A., Bauer, M., Pinotsis, D., & Friston, K. J. (2016). Dynamic causal modelling of eye movements during pursuit: Confirming precision-encoding in V1 using MEG. *NeuroImage*, 132, 175–189.
- Ao, P. (2008). Emerging of Stochastic Dynamical Equalities and Steady State Thermodynamics. *Commun. Theor. Phys.* (Beijing, China), 49, 1073-1090.
- Arnal, L. H., Wyart, V., & Giraud, A.-L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, 14(6), 797–801.
- Badcock, P. B. (2012). Evolutionary systems theory: A unifying meta-theory of psychological science. *Review of General Psychology: Journal of Division 1, of the American Psychological Association*, 16(1), 10–23.
- Badcock, P. B., Davey, C. G., Whittle, S., Allen, N. B., & Friston, K. J. (2017). The Depressed Brain: An Evolutionary Systems Theory. *Trends in Cognitive Sciences*, 21(3), 182–194.
- Badcock, P. B., Ploeger, A., & Allen, N. B. (2016). After phrenology: Time for a paradigm shift in cognitive science. *The Behavioral and Brain Sciences*, 39, e121.
- Berger, P. L. & T. Luckmann. (1966). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Anchor Books.
- Branco, T., Clark, B. A., & Häusser, M. (2010). Dendritic discrimination of temporal input sequences in cortical neurons. *Science*, 329(5999), 1671–1675.
- Bruineberg, J. (2018). Anticipating affordances: Intentionality in self-organizing brain-body-environment systems (Doctoral dissertation). Retrieved from UvA-DARE.
- Bruineberg, J., & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neuroscience*, 8, 599.
- Campbell, J. O. (2016b). Universal Darwinism As a Process of Bayesian Inference. *Frontiers in Systems Neuroscience*, 10, 49.
- Chemero, A. (2009). *Radical embodied cognition*. Cambridge, MA: MIT Press.
- Clark, A. (2015). *Surfing uncertainty: prediction, action, and the embodied mind*. New York, N.Y.: Oxford University Press.
- Colombo, M. (2014). Explaining social norm compliance. A plea for neural representations. *Phenomenol. Cogn. Sci.* 13, 217–238
- Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, 11(2), 127–138.
- Constant, A., Bervoets, J., Hens, K., & Van de Cruys, S. (2018). Precise Worlds for Certain Minds: An Ecological Perspective on the Relational Self in Autism. *Topoi. An International Review of Philosophy*, 1–13.
- Constant, A., Ramstead, M. J. D., Veissière, S. P. L., Campbell, J. O., & Friston, K. J. (2018). A variational approach to niche construction. *Journal of the Royal Society, Interface*, 15(141).
- Dawkins, R. (1976). *The Selfish Gene*, New York: Oxford University Press.
- Edelman, G. M. (1987). *The Theory of Neuronal Group Selection*. New York: Basic Books.
- Engel, A. K., Friston, K. J., & Kragic, D. (2016). *The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science*. MIT Press.
- Fragaszy, D. M. (2011). Community Resources for Learning: How Capuchin Monkeys Construct Technical Traditions. *Biological Theory*, 6(3), 231–240. <http://doi.org/10.1007/s13752-012-0032-8>
- Fragaszy, D. M., Eshchar, Y., Visalberghi, E., Resende, B., Laity, K., & Izar, P. (2017). Synchronized practice helps bearded capuchin monkeys learn to extend attention while learning a tradition. *Proceedings of the National Academy of Sciences*, 114(30), 7798–7805.
- Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, 11(2), 127–138.

- Friston, K. J. (2013). Life as we know it. *Journal of the Royal Society, Interface / the Royal Society*, 10(86), 20130475.
- Friston, K., & Ao, P. (2012). Free-energy, value and attractors. *Computational and mathematical methods in medicine*, 937860.
- Friston, K. J., Daunizeau, J., & Kiebel, S. J. (2009). Reinforcement learning or active inference? *PloS One*, 4(7), e6421.
- Friston, K. J., & Frith, C. D. (2015). Active inference, communication and hermeneutics. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 68, 129–43.
- Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, 104(1-2), 137–160.
- Friston, K., Thornton, C., & Clark, A. (2012). Free-energy minimization and the dark-room problem. *Frontiers in Psychology*, 3, 130.
- Gibson, J. J. (1979). *The ecological approach to visual perception: classic edition*. Psychology Press.
- Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge University Press.
- Gulledge, A. T., Kampa, B. M., & Stuart, G. J. (2005). Synaptic integration in dendritic trees. *Journal of Neurobiology*, 64(1), 75–90.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7(1), 1–16.
- Hickok, G. (2013). Predictive coding? Yes, but from what source? *The Behavioral and Brain Sciences*, 36(4), 358.
- Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.
- Huygens, C. (1673). *Horologium oscillatorium*. France: Parisiis.
- Kiebel, S. J., Daunizeau, J., & Friston, K. J. (2008). A hierarchy of time-scales and the brain. *PLoS Computational Biology*, 4(11), e1000209.
- Kiebel, S. J., & Friston, K. J. (2011). Free energy and dendritic self-organization. *Frontiers in Systems Neuroscience*, 5, 80.
- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, 8(3), 159–166.
- Kirchhoff, M. (2017). Predictive brains and embodied, enactive cognition: an introduction to the special issue. *Synthese*, 1–12.
- Kirchhoff, M., Parr, T., Palacios, E., Friston, K., & Kiverstein, J. (2018). The Markov blankets of life: autonomy, active inference and the free energy principle. *Journal of the Royal Society, Interface / the Royal Society*, 15(138). <https://doi.org/10.1098/rsif.2017.0792>
- Kirchhoff, M.D. (2017). Predictive processing, perceiving and imagining: Is to perceive to imagine, or something close to it? *Philosophical Studies*, 1-17, doi: 10.1007/s11098-017-0891-8.
- Kirchhoff, M.D. (2015). Species of realization and the Free Energy Principle. *The Australasian Journal of Philosophy*, 93(4), 706-723.
- Lendvai, B., Stern, E. A., Chen, B., and Svoboda, K. (2000). Experience-dependent plasticity of dendritic spines in the developing rat barrel cortex in vivo. *Nature* 404, 876–881.
- Manneville, P. (1995). Dissipative Structures and Weak Turbulence. *Springer Lecture Notes in Physics*, 457, 257–272.
- Maynard Smith, J. (1964). Group selection and kin selection. *Nature*, 201(4924), 1145-1147.
- McKinley, J. (2015). Critical Argument and Writer Identity: Social Constructivism as a Theoretical Framework for EFL Academic Writing. *Critical Inquiry in Language Studies*, 12(3), 184–207.
- Naiman, R. J., Johnston, C. A., & Kelley, J. C. (1988). Alteration of North American Streams by Beaver The structure and dynamics of streams are changing as beaver recolonize their historic habitat. *Bioscience*, 38(11), 753–762.
- Noë, A. (2004). *Action in Perception*. MIT Press.
- Odling-Smee, F. J., & Laland, K. N. (2000). Niche Construction and Gene-Culture Coevolution: An Evolutionary Basis for the Human. In T. N. S. Tonneau F. (Ed.),

- Perspectives in Ethology* (Vol. Sciences Perspectives in Ethology, 13). Boston, MA: Springer.
- Odling-Smee, J., Erwin, D. H., Palkovacs, E. P., Feldman, M. W., & Laland, K. N. (2013). Niche construction theory: a practical guide for ecologists. *The Quarterly Review of Biology*, 88(1), 4–28.
- Odling-Smee, J., Laland, K. N., & Feldman, M. W. (2003). *Niche Construction: The Neglected Process in Evolution*. Princeton University Press.
- Palacios, E. R., Razi, A., Parr, T., Kirchhoff, M., & Friston, K. (2017, November 30). *Biological Self-organisation and Markov blankets*. *bioRxiv*. bioRxiv.
- Orgel, L. E., & Crick, F. H. C. (1980). Selfish DNA: the ultimate parasite. *Nature*, 284(5757), 604–607.
- Pearl, J. (1988). Probabilistic reasoning in intelligent systems: Networks of plausible inference. San Mateo, CA: Morgan Kaufmann.
- Ramstead, M. J. D., Badcock, P. B., & Friston, K. J. (2017). Answering Schrödinger’s question: A free-energy formulation. *Physics of Life Reviews*. <https://doi.org/10.1016/j.plrev.2017.09.001>
- Ramstead, M. J. D., Badcock, P. B., & Friston, K. J. (2018). Variational neuroethology: Answering further questions: Reply to comments on “Answering Schrödinger’s question: A free-energy formulation.” *Physics of Life Reviews*, 24, 59–66.
- Ramstead, M. J. D., Veissière, S. P. L., & Kirmayer, L. J. (2016). Cultural affordances: scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in Psychology*, 7, 1090.
- Seifert, U. (2012). Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep Prog Phys*, 75(12), 126001. doi: 10.1088/0034-4885/75/12/126001
- Sengupta, B., Stemmler, M. B., & Friston, K. J. (2013). Information and Efficiency in the Nervous System—A Synthesis. *PLoS Computational Biology*, 9(7).
- Sengupta, B., Tozzi, A., Cooray, G. K., Douglas, P. K., & Friston, K. J. (2016). Towards a Neuronal Gauge Theory. *PLoS Biology*, 14(3), e1002400.
- Stotz, K. (2017). Why developmental niche construction is not selective niche construction: and why it matters. *Interface Focus*, 7(5), 20160157.
- Sun, Y., Gomez, F., #252, & Schmidhuber, r. (2011). Planning to be surprised: optimal Bayesian exploration in dynamic environments. Paper presented at the Proceedings of the 4th international conference on Artificial general intelligence, Mountain View, CA.
- Thompson, E. (2007). *Mind in life: biology, phenomenology, and the sciences of mind*. Cambridge, MA: Harvard University Press.
- Torben-Nielsen, B., & Stiefel, K. M. (2009). Systematic mapping between dendritic function and structure. *Network*, 20(2), 69–105.
- van Dijk, L., Withagen, R., & Bongers, R. M. (2015). Information without content: A Gibsonian reply to enactivists’ worries. *Cognition*, 134, 210–214.
- Varela, F. G., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: the organization of living systems, its characterization and a model. *Currents in Modern Biology*, 5(4), 187–96.
- Varela, F. J., Thompson, E., & Rosch, E. (2017). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.
- Weiss, L., Brandl, P., & Frynta, D. (2015). Fear reactions to snakes in naïve mouse lemurs and pig-tailed macaques. *Primates*, 56(3), 279–284.
- Yi, S., Wierstra, D., Schaul, T., & Schmidhuber, J. (2009). Stochastic search using the natural gradient. In *Proceedings of the 26th Annual International Conference on Machine Learning - ICML ’09* (pp. 1–8). New York, New York, USA: ACM Press.