This is a preprint accepted for publication in Synthese. Please cite the published version.

Between pebbles and organisms:

Weaving autonomy into the Markov blanket

Thomas van Es[1]* & Michael D. Kirchhoff[2]

Affiliations:
[1]Centre for Philosophical Psychology, Department of Philosophy, Universiteit Antwerpen
[2]School of Liberal Arts, Faculty of Arts, Social Sciences and the Humanities, University of Wollongong, Wollongong, Australia.
* Corresponding author

Acknowledgements: Kirchhoff's work was supported by an Australian Research Council Discovery Project "Mind in Skilled Performance" (DP170102987). Van Es's work was supported by the Research Foundation Flanders (Grant No. 1124818N). We would like to thank Mel Andrews, Mads Julian Dengsø and two anonymous reviewers for comments on a previous draft of this paper.

Abstract: The free energy principle (FEP) is sometimes put forward as accounting for biological self-organization and cognition. It states that for a system to maintain non-equilibrium steady-state with its environment it can be described as minimising its free energy. It is said to be entirely scale-free, applying to anything from particles to organisms, and interactive machines, spanning from the abiotic to the biotic. Because the FEP is so general in its application, one might wonder whether this framework can capture *anything specific* to biology. We take steps to correct for this here. We first explicate the worry, taking pebbles as examples of an abiotic system, and then discuss to what extent the FEP can distinguish its dynamics from an organism's. We articulate the notion of 'autonomy as precarious operational closure' from the enactive literature, and investigate how it can be unpacked within the FEP. This enables the FEP to delineate between the abiotic and the biotic; avoiding the pebble worry that keeps it out of touch with the living systems we encounter in the world.

Keywords: Free energy principle; Markov blanket; autonomy; operational closure; biology; cognition; the pebble challenge; unification

1 Introduction

The free energy principle (FEP) is a *principle first* approach to what it takes for a system to exist. Rather than empirical investigation, the FEP starts from a mathematical *principle* that a system is thought to conform to if it exists. Indeed, FEP researchers seek to provide a general theory unifying biology and cognitive science formulated almost entirely from mathematical principles in physics and information theory (see e.g., Friston 2010 2013; Hohwy 2020; Kirchhoff et al. 2018; Linson et al. 2018; Ramstead, Kirchhoff, Friston 2019). The ambition is to secure a definition of existence by appealing to constructs in physics and information theory, and then employing those constructs to derive a principle of self-organization and cognition (Friston 2019; Hesp et al. 2019). In a nutshell, the FEP states that a system that maintains non-equilibrium steady-state (NESS) with its environment can necessarily be cast as minimising free energy.[1] This particular observation can consequently be exploited to show a wide variety of interesting relations to hold between a NESS system and its environment.

Yet the FEP's mathematical toolkit is not only applicable to living systems. It is said to be entirely *scale-free* in its applicability. That is, it is intended to apply to any system able to maintain its organisation despite tendencies towards disorder: from chemotaxis in cells (Friston 2013; Auletta 2013), neuronal signalling in brains (Friston et al. 2017; Parr & Friston 2019), tropism in plants (Calvo & Friston 2017), synchronised singing in birds (Frith & Friston 2015) to decision-making and planning in mammals (Daunizeau et al. 2010; Friston 2013; Williams 2018). It has also been applied to model adaptive fitness over evolutionary timescales by casting evolution in terms of Bayesian model optimisation and selection (Campbell 2016; Hesp et al. 2019). However, this widespread applicability of the FEP can be taken as a fault, rather than an advantage.

Indeed, there is a general concern about the FEP's ability to speak to the essential organizational dynamics of biology, because it can seem utterly disconnected from biology. More specifically, the FEP is sometimes considered incapable of uniquely addressing the organisational dynamics of living systems (van Es 2020; Colombo and Wright 2018). Because the FEP implies an entirely scale-free dynamics in which *any* self-organising NESS system can be cast in terms of self-evidencing, some worry that this particular view cannot

---

[1] The term 'non-equilibrium steady-state' refers to *self-sustaining* processes in a system requiring input and output to avoid relaxing into *thermodynamic equilibrium* (= systemic decay/death). It is important to mention here that the notion 'steady-state' in non-equilibrium systems is an approximation to some specified duration of time - e.g., circadian rhythms over a 24 hour clock cycle or the homeostatic processes involved in maintaining on average and over time a specific body temperature. So strictly speaking, biological systems are not in steady states; rather, to say that a system is in a steady-state, *X*, at a particular time, is effectively to say that the probability density over the system's states during some period of time was *X*.

capture the specific details of biological organisation that is of interest to the biological sciences. If true, this undercuts the grand unifying ambitions of many FEP researchers.

We address this worry here. We start by rehearsing the basic tenets of the FEP, with particular focus on the Markov blanket formalism and how it relates to Bayesian inference (sect. 2). We proceed to explicate the aforementioned worry by considering the application of the FEP formalism to a pebble and discuss how the FEP seems to fall short in delivering the tools to distinguish pebbles from organisms (sect. 3). *Prima facie,* its scale-free applicability makes it seem like it is unable to carve any interesting joints between the abiotic and the biotic, which would hinder the prospect of a FEP biology. Kirchhoff et al. (2018) make an initial attempt to address this problem, suggesting that *autonomy* is what distinguishes living from non-living systems. The overarching claim there is that autonomy is the capacity of a system to modulate its exchange with its environment. Here we supplement this initial treatment. We first look at 'autonomy' from an enactive viewpoint (sect. 4). We then sketch the contours of how the notion of 'autonomy' from the enactive literature could be emulated with the tools available to the FEP formalisms. This allows us to understand what constitutes an autonomous system rather than merely using the notion of autonomy as a mark by which to delineate life from no-life (sect. 5).

2 Markov blankets, free energy and Bayesian inference

The FEP speaks to what characteristics a system must exhibit for it to exist (Friston 2013). Its basic premise is that any random dynamical system "that possesses a Markov blanket will appear to actively maintain its structural and dynamical integrity" (Friston 2013, p. 2).

A Markov blanket is a statistical separation of states that is applicable to any thing that exists (Hipólito 2020). It is a set of blanket states that separates a system's internal states from external states (Pearl, 1988; Beal 2003). The blanket states shield (in a statistical sense) internal from external states, and vice versa. They can be partitioned into sensory states and active states. Sensory states capture the influence of external states on internal states. Active states capture the influence of internal states on external states. Intuitively, any thing can be separated statistically from that which it is not (Palacios et al. 2020).

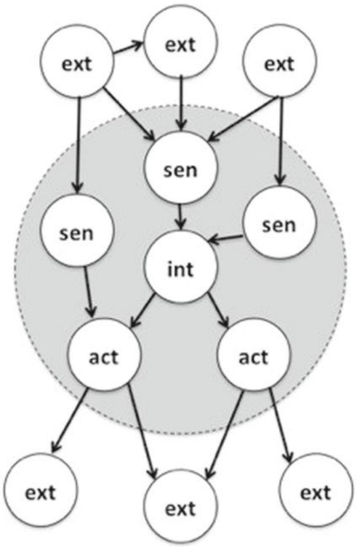

Figure 1 is a schematic representation of a Markov blanketed system. The gray circle delineates the Markov blanketed system that separates internal states (int) from external states (ext). The blanket states, sensory states (sen) and active states (act) are displayed as surrounding the internal states. The arrows depict modes of influence. External states influence only one another or sensory states, while influenced only by active states or one another. Internal states are influenced only by sensory states, influencing only active states. In terms of modes of influence, internal states are separated from external states. (figure taken from Bruineberg, et al. 2018).

In this statistical formulation, the separation between internal and external states implies that these states are *conditionally independent*, given the states that comprise the Markov blanket. If we want to figure out the *external* states and we *know* the values of the *blanket* states, *knowing* the values of the *internal* states *will not offer additional predictive value*, and vice-versa. This is so by definition, because the blanket states already capture any possible influence the internal states could have on the external states. A brief example may clarify this. Say you observe that it is cold. This could be either due to an open window or to an air conditioning system that is set too strong. If you would observe that, say, the air conditioning is set excessively high, the observation that it is cold now does not offer further information with regards to whether or not the window is open. That is, in this case, the observed cold and the open window are conditionally independent, given that the air conditioning is on blast (Kirchhoff and Kiverstein 2019; Beal 2003). In terms of the Markov blanket formalism, the observed cold could be cast as the internal states, the state of the window could be formalized as the external states with the states of the air conditioning serving as the blanket states. This example is important because it indicates the widespread applicability of the formalism. Indeed, it is not necessarily obvious to associate the boundaries induced by the Markov blanket with physical boundaries, though it does seem to lend itself well to this particular application. We should nonetheless remain wary about overstating the implications of this *statistical* partitioning of states when considering its application onto other systems (van Es 2019, 2020).

Now that it is clear what a Markov blanket is, we can delve into its relation with free energy minimisation and Bayesian inference. This is a technical story. According to the Second Law of thermodynamics, the *entropy* of any closed system increases indefinitely over time. Any system that exists, or any Markov blanketed system that retains its structural integrity over time, seems to temporarily slow down the increase of entropy for as long as it remains intact (Friston 2012, 2013, 2019; Schrödinger, 1944). Of course, any such 'resistance' is only temporary, as entropy increases upon disintegration, which, in the case of biotic systems, means death.

For any such system, you can establish a multi-dimensional state space with as many dimensions as there are variables represented in the state space. Each point in the state space corresponds to a unique intersection of values for each variable. In this state space, you can mark a bound of states within which a system can remain intact, outside of which it cannot (Friston, 2012, 2013). For as long as the system remains intact, the system will continuously 'revisit' the states within this bound. This is so by definition, as we define the bound by the range of values within the system remains intact. With regards to organisms, the viable bound differs per species: humans remain intact under quite different circumstances than fish do, for example. Insofar as this bound counts as a description of the states in which the system can be found when alive, it is also considered to be a mathematical description of a phenotype (Friston 2013; Kirchhoff et al 2018).[2] On average and over time, any living system is thus likely to be found within the bound of viable states and unlikely to be found outside of it. That is, we may *expect* a system to be within a bound of states that it typically remains within *on average* (Friston, Wiese, Hobson 2020). This implies a probability distribution that can be laid over the state space so that each state is assigned a probability value (Ramstead et al. 2019; Corcoran et al. 2020; Friston 2013). At any given time the system is encountered, it is highly likely to occupy a state within the viable bound, and highly unlikely to occupy a state outside of this. This means that states within the bound are considered high-probability states, whereas states outside of it are considered low-probability states.

Furthemore, if a system's *internal* states remain within a particular range, this must mean that the *influences* on those states are similarly bounded. An example should clarify this. Consider an egg and spoon race. An egg-and-spoon runner will need to ensure that the influence on the egg of their running the race remains within certain bounds, lest the egg move out of the spoon and break. Let us apply the Markov blanket partitioning method. We shall take the internal states here to be the egg's, and the influences it receives via the spoon shall be the sensory states, the runner is here the environment impacting on the spoon and comprises the

---

[2] See Colombo and Wright (2018) for criticism on the viability of this application onto an organismic system.

external states. This means that an egg-and-spoon runner can be cast as keeping a *tight bound on the sensory states* of the Markov blanketed egg for as long as it remains in the spoon.

As it is for the egg-and-spoon runner, so it is for *any* system that remains intact over time. Relative to the viable bound of the internal states of the system, then, we can also establish *a state space for the sensory states* within which the system can remain intact, outside of which it cannot. Here too, we can determine a probability distribution where states within the bound are ascribed high probability, those outside of it are ascribed low probability. This is a probability distribution *over external states*, as it relates to the influences on the internal states *by the external states*. In other words, it defines the possible external states that there could be relative to the internal states, given that the internal states remain within the viable bound. Of interest here is that the internal states themselves provide all we need (the ‘sufficient statistics’) to compute the probability distribution over the external states. As such, by knowing the viable bound of the internal states, we can compute the viable bound of the system’s sensory states.

Further, in Bayesian probabilistic theory, *surprise* is a quantity defined as the *improbability* of a particular state (Shannon, 1948). If the surprise of sensory states (or ‘sensory surprise’, not to be confused with agent-level surprise with regards to an unexpected sensation) is high, the sensory states currently occupy a low probability area in the state space. As low-probability states are those that endanger the system’s structural integrity, surprise is kept low, or minimized, as long as the system remains intact. However, sensory surprise is a probabilistic measure of sensory states. The entire state space of sensory states includes all possible modes of influence the external states could possibly exert on the internal states. This is, in principle, an infinite set. Computing sensory surprise directly is thus intractable (Friston 2009).

This is where (variational) *free energy* comes in. Free energy, in the statistical usage of the term, is a functional of the internal and sensory states a system is in.[3] In this case, free energy is thus, more specifically, the function of a function of the sensory states that is parameterized by the internal states. Because of this, the value of free energy limits the possible values of the internal and sensory states. To see why, consider a solution to a simple summation problem in arithmetics, say it’s 15, and the terms of the equation are non-negative. This means that none of the terms of the problem can exceed the value of 15. Minimizing the value of free energy, then, minimizes an upper bound on the probability of sensory states. This ensures that sensory states remain in high-probability areas in the state space, which in turn implies that sensory surprise is minimized. Minimizing free energy can thus be seen as approximately minimizing the otherwise intractable value of sensory surprise (Friston and Stephan 2007). Moreover, as free energy is a function of only the internal and sensory states, it is in principle computable (Kiebel, Daunizeau and Friston 2008; Friston and Ao 2012).

---

[3] A functional is a function of a function.

In Bayesian probability theory, *negative surprise* is equivalent to Bayesian model evidence. *Minimizing* surprise thus *maximizes* Bayesian model evidence. The process by which Bayesian model evidence can be maximized is called *Bayesian inference*, sometimes referred to as self-evidencing (Friston, Killner, Harrison 2006; Hohwy 2016). Bayesian inference then refers to the particular way a probability distribution needs to be updated in light of new evidence (Beal 2003). Bayesian inference describes the permissible 'moves' one can make in the formal system of Bayesian probability theory. We can now see that for any system to remain intact over time, its entropy needs to be minimized on average over time, which means expected free energy needs to be minimized, which in turn implies the minimization of sensory surprise, which is done by way of a formal operation called Bayesian inference.

The above story is employed in the FEP as a mathematical description of the homeostatic processes of biotic systems (Friston 2013). This works, very roughly, as follows. In the Markov blanket formalism, the Markov blanket is thought to carve out ontological joints: the internal states map onto the organism itself, and the external states map onto the environment (Kirchhoff and Kiverstein 2019). The partitioning blanket states map onto the organism's modes of interaction so that sensory states are associated with sensory receptor activity, and active states are associated with the system's influence on its environment, such as action. It remains a current debate to what extent this application of the Markov blanket should be taken literally or instrumentally (van Es 2020; Bruineberg et al. 2020; Hohwy 2016). In this paper, we will remain neutral in this debate, and instead explore only what can be done within the formalism, regardless of how it may or may not be implemented in any real system.

In a realist interpretation, to 'engage' in Bayesian inference is considered a fundamental aspect of life, as without it, the organism would go outside of its viable bounds. This is called *active inference*, and is thought to account for both action and perception by the same guiding principle (Friston 2013). The probability distributions are embodied and/or encoded by the organism (and/or the brain). They are to be manipulated, updated and leveraged by the organism. Through active inference, the organism updates the probability distributions in the face of newfound evidence, and uses this to infer action policies for its interaction with the world. Long term activities are thought to require counterfactual inference, which is associated with the minimization of *expected free energy* or free energy on average over time (Corcoran et al. 2020). Rather than updating the probability distribution to remain within its viable bounds, this should be seen as the inference of a possible trajectory through the state space conditioned on bodily movement. This allows the organism to adapt to environmental fluctuations. After all, the distribution of states within which an organism can remain alive cannot be simply 'updated' when confronted with an environment likely to push the system outside of viable bounds. Active inference thus plays a central role in the realist FEP story of biological systems.[4]

3 Pebble meets Markov blanket

One person's meat is another person's poison: the scale-free applicability of the FEP's Markov blanket formalism may be taken as a vice, rather than a virtue. In this section we take up a specific challenge to the FEP that flows from what seems like an overly generous application of the FEP formalism to a wide variety of phenomena: the pebble challenge. It challenges the FEP's ambitions to describe the organizational dynamics of life precisely because its mathematical formalisms apply equally well to pebbles, and other abiotic systems as they do to biotic ones. One might therefore worry that the FEP fails to say anything specific about biology, unless characteristics we take to be specific to biology are not so specific at all. We describe this challenge in more detail now.

Friston & Stephan (2007) anticipates this kind of challenge to the FEP. They ask, "What is the difference between a plant [a biotic system] and a stone [an abiotic system]?" (2007, p. 422) They say that the plant "is an open non-equilibrium system, exchanging matter and energy with the environment, whereas the stone is an open system that is largely at equilibrium" (2007, p. 422). There is something to this initial observation. Plants are open systems, i.e., energy and mass can flow between the system and its surroundings. The same, of course, can be said of stones as environmental forces impinge on their surface area, and their own existence influences their environment by, say, releasing heat during the day, or altering pathways for organisms (Olivotos & Economou-Eliopoulos 2016). At first glance, it thus seems that the FEP applies in the same way to stones, plants and humans.

The FEP (as we saw above) starts from the simple observation "that *for something to exist* it must possess (internal or intrinsic) states that can be separated *statistically* from (external or extrinsic) states that do not constitute the thing" (Friston 2019, p. 4, emphases added). This Markov blanket formulation would apply to a pebble as follows. The Markov blanket defines the conditional independencies between two sets of states: the system and the environment. Pebbles are composed of minerals with different properties, lattice structure, hardness and cleavage. We can associate these variables as the internal states comprising the system. On shingle beaches, the second set of states (the environment) would be other pebbles, and so on. In rivers, the water could be cast as the external states. As seen in Section 2, it is possible to cast a spatial boundary for anything that exists in terms of a Markov blanket (Friston 2013). The pebble has a clear boundary separating internal states and external states. The sensory

---

[4] The extent to which this story should be taken in a realist sense so that each biotic system literally performs advanced statistical operations, or in an instrumentalist sense so that each biotic system's interactional dynamics merely correspond to (or 'instantiate') the dynamics described in Bayesian inference is still debated (van Es 2020; see Ramstead, Kirchhoff, Friston 2019; Corcoran et al. 2020). A discussion of this debate is outside the scope of this paper, and it is unnecessary for our current purposes.

states of a pebble can be associated with the effects of external causes of its boundary - stressors such as pressure, temperature and so on. Its active states would correspond to how the pebble effects external states - e.g., via release of heat back into the environment. The Markov partitioning rule governing the relation between states dictates that external states act on sensory states, which influence, but are not themselves influenced by internal states. Internal states couple back to external states, via active states, which are not influenced by external states (Palacios et al. 2020). Given that the Markov blanket formulation for a pebble is possible, it follows that internal pebble states are conditionally independent of external states in virtue of the Markov blanket states.[5]

What does this mean for our FEP analysis of the pebble, given what we have seen in Section 2? Under the FEP, the mere presence of a Markov blanket implies that internal states can be understood as if they minimise the free energy over the states that make up their Markov blanket. Technically, since minimising free energy is the same as performing approximate Bayesian inference, it follows that one can associate the internal pebble states (and its blanket states) with Bayesian inference. As such, it seems that if (1) anything that exists over time can be described in terms of a Markov blanket which implies that expected free energy is minimized by way of Bayesian inference, and (2) pebbles exist, then (3) pebbles can be described as having a Markov blanket, whose dynamics will appear as though they minimize free energy by way of Bayesian inference. The formalisms of the FEP that we employed here therefore seem too general to distinguish between pebbles and organisms. Below we will discuss what is needed for a formalism to properly address autonomy in Section 4. In Section 5 we will see how FEP's toolkit can be leveraged to make a headway in providing a principled distinction between pebbles and organisms.

4 Autonomy meets pebble

The pebble challenge need not be a knockdown argument against the ambitions of the FEP to address biology and cognitive science. Here we consider a possible reply to it. Our agenda will be to introduce the notion of *autonomy* from enactive philosophy of cognitive science.[6]

---

[5] We could, for example, determine the surface molecules of the pebble to be sensory states, adjacent molecules to be active states, and the remainder of the pebble's molecules to be internal states, with the environment cast as external states. The molecules we cast as active states are then shielded from influence of the external states, while still able to influence the external states, though vicariously through sensory states. Of course, a pebble is merely an example and this could apply to many abiotic systems. Thanks to an anonymous reviewer for pointing this out.

[6] Autonomy is a central theoretical construct of the enactive approach to life and mind (Varela, 1979; Varela et al., 1991; Thompson 2007; Di Paolo & Thompson 2014; Di Paolo et al. 2017). Enactivism is a theoretical framework with roots in theoretical biology, dynamic systems theory, and phenomenology. In enactivism, the notion of autonomy as operational closure has received special attention in attempting to unearth the self-organisational dynamics essential to life. Yet the literature so far has fallen short of construing operational closure in terms of the FEP's conceptual toolkit. Here we will make a first attempt at conceiving of an operationally closed system as being composed of a network of Markov blanketed systems that stand in a mutually enabling relation to one another.

Kirchhoff et al. (2018) appeal to this notion in order to distinguish between *mere active inference* and *adaptive active inference*. The former can be shown to apply to abiotic systems such as pebbles (from above) and the generalised synchrony induced in coupled pendulum dynamics. Adaptive active inference is introduced to make sense of the idea that living organisms are able to actively change or modulate their sensorimotor coupling to their environment - which is needed to actively monitor and predict changes to perturbations that challenge homeostatic variables, which may, sometimes, go out of bounds. However, the modulation of sensorimotor coupling is merely a (contingent) feature of an autonomous system. Operational closure and precariousness jointly define autonomy. We build on Kirchhoff et al's (2018; see also Kirchhoff & Froese 2017) argument by showing how autonomy is underwritten by the concepts of operational closure and precariousness (cf. Di Paolo & Thompson 2014).

4.1 Operational closure and precariousness

Operational closure is central to the conceptualisation of autonomy (Di Paolo & Thompson 2014). It is characterized as a form of *organization* in the sense that it specifies the particular way any system's component parts are organized in relation to one another. By specifying the organized 'unity' (the system) via this formalism, we also implicitly define its environment. Furthermore, by defining the system and its environment, we also specify the boundary through which the system interacts with its environment (Beer 2004, 2014; Maturana and Varela, 1980).

A system is operationally closed if the processes that make up the system constitute what is known as a self-enabling network. This means that each of the network's processes enables and is enabled by at least one other process in the network. It is empirically possible to determine whether any particular system is operationally closed by mapping out the causal processes relevant for the system and how they relate to one another. In particular, one must look for *enabling* relations. Any one process is said to enable another process if its continuation is partly or wholly constitutive of the enabled process. To explain how this works, it may help to look at the diagram of an operationally closed system (Figure 2 below). In this toy system, we distinguish five component processes: A, B, C, D, and E represented as nodes in the figure. The arrows between them represent enabling relations, so that A can be seen to enable process B. Following the arrows, we can identify a closed loop in the enabling relations pertaining to processes A, B, and C. This means that the continuation of A enables the continuation of B, which enables the continuation of C, which comes full circle and enables the continuation of process A: the ABC network is thus *self-enabling*. But what about processes D and E? E can be seen to enable process A, yet remains outside of the network as it is not enabled by a process in the network. D, on the other hand, *is* enabled by a process in the network, but doesn't loop back and enable a process in the network itself. This is why ABC can be identified as a self-enabling network, while D and E fall outside the boat.

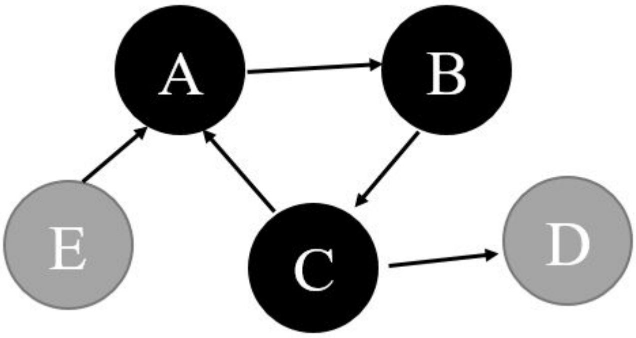

Figure 2, a diagram displaying an operationally closed network of enabling relations. Each node in the figure represents a process in the network, and each arrow represents an enabling relation. The operationally closed network is marked by the black nodes; processes outside the operationally closed system are marked in grey. Each node that is part of the operationally closed network is marked by having at least one outgoing and one incoming arrow from another node in the operationally closed network as is described in-text (inspired by Di Paolo and Thompson 2014).

Precariousness signifies a natural inclination to decline. In Figure 2 above, for example, process A is precarious if it would cease were it not enabled by E and C. It may be that not each enabling process is *per se* necessary or sufficient in enabling. If A is precarious, this does mean however, that jointly, its enablers are both necessary and sufficient for the continuation of A. As each node in the network is precarious, the network itself is too. This is crucial for the notion of autonomy in the enactive approach (Di Paolo 2005).

A paradigmatic case that displays operational closure and precariousness is a single cell. A cell is constituted by a complex network of interrelated causal processes, but, for didactic purposes, we distinguish three. The first process comprises the metabolic network. The second process is the membrane-generation of the cell that separates the network from the environment. The third process consists of the active regulation of matter and energy exchanges of the cell, via the membrane-induced barrier, with its external environment. By way of this third process, the system can absorb nutrients from and expel wastes into its environment to continue its metabolism, looping back into process one.

The metabolic network, process 1, can be divided into subprocesses. A central aspect of metabolism is the production of enzymes, which exhibits a form of closure in itself. Enzymes are precarious. As such, when particular enzymes need to be produced, this occurs “in metabolic pathways helped by other enzymes, which in turn are produced with the participation of other ones … in a *recursive* way” (Mossio and Moreno 2010, p. 278, emphasis added). That is, the metabolic network in itself can be said to be “enzymatically closed” (Mossio and Moreno 2010, p. 278). This production network enables process 2: the generation of a membrane that separates the network from its environment. This semipermeable barrier is necessary for the system to actively regulate its exchanges with the environment. It both allows the system to take in matter and energy from the environment, and protect its internal network from external perturbation of the metabolism (Ruiz-Mirazo and Mavelli 2008; Thompson 2007). The exchange with the environment enabled by the barrier's separation is process 3. The limited openness is exploited to allow for the absorption of nutrients from the environment which can stimulate the maintenance of the membrane itself, but also "contribute to the production of an 'energy currency'" (Ruiz-Mirazo and Mavelli 2008, 376; Skulachev, 1992). Via trans-membrane mechanisms, this 'currency' is cashed out in internal metabolic reactions, transformed to serve as energy resources to maintain and actively regulate its boundary conditions (Ruiz-Mirazo, Mavelli 2008). This is to say that process 3 loops back into enabling process 1 and 2. These enabling relations are visualized in Figure 3 below. Here we can see that operational closure and precariousness jointly correctly marks a cell as an autonomous system.

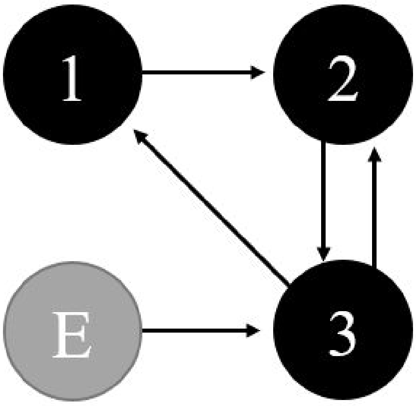

Figure 3 illustrates the simplified process network relevant to a single cell. Process 1, which captures the metabolic network, is represented by the 1 in the top-left. Process 2, membrane-generation, is represented by 2 in the top-right. Process 3, the active regulation of matter and energy exchanges with the environment is represented by 3 in the bottom-right. The environment is represented by the E in the bottom-left. The arrows between the represented processes stand for enabling relations as described above. We see that 1, 2 and 3 form a self-enabling network as per the definition above. Each process in the network enables and is enabled by at least one other process in the network. 1 enables 2 and is enabled by 3. 2 enables 3 and is enabled by 1 as well as 3. 3 enables both 1 and 2, and is enabled by 2 and E. The network here described thus represents an operationally closed system.

4.2 Autonomy and the pebble

A pebble is not autonomous. Given that autonomy is intended to solve the pebble challenge, it is important to subject the pebble to the same analysis: is a pebble operationally closed and precarious? If not, this indicates that autonomy as used here is an adequate concept to distinguish between abiotic and biotic systems. We distinguish four causal processes that are relevant to the formation and maintenance of the pebble's structural integrity on a shingle beach, two of which are directly considered to be determinants of a pebble's shape and size: particle abrasion and particle transport. These two processes may be more or less relevant depending on the particular geological location (Domokos and Gibbons 2012; see also Landon, 1930; Kuenen, 1964; Carr, 1969; Bluck, 1967). Particle transport refers to the transport of the pebble by the river. Particle abrasion refers to the collusion with other pebbles (and other materials) that occurs primarily during particle transport. The remaining two processes are the fluid flows of the river and the environment that consists of abraders of a hard enough consistency to allow for particle abrasion.

The four processes in the network are thus: fluid flows (A), environmental abraders (B), particle abrasion (C) and particle transport (D). Fluid flows enable particle transport, and can reasonably be considered to enable particle abrasion too. Assuming there are no other moving objects in the river, the pebble will be unlikely to move from its location and is thus unlikely to be abraded by other materials, if it is not swept anywhere by the fluid flows. Environmental abraders only enable particle abrasion. Particle abrasion in itself does not enable any other process in the network. Particle transport only enables particle abrasion. This means that fluid flows only enable other processes, but are themselves not enabled by any other process in the network. The enabling relations are specified in Figure 4 below. This means that A cannot be part of a self-enabling network. Environmental abraders only enable particle abrasion, and are not themselves enabled by other processes in the network and thus B suffers the same fate as A. C, particle abrasion, is enabled by all other processes in the network, but does not actually enable any other process, and can also not figure in a self-enabling network. Process D, particle transport, is the only process that is both enabled by and enables another process in the network, being enabled by fluid flows, enabling particle abrasion. This enabling chain, however, never loops back into enabling the continuation of particle transport. As such, Process D too cannot be part of a self-enabling network. Summing it up, there is no self-enabling network to be found in the processual network surrounding pebbles. This means that, under the operational closure formalism, pebbles are not marked as autonomous.[7]

---

[7] Our treatment of the pebble case may seem disanalogous with our treatment of the cell case. The discussion of the cell case treated a few important *internal* processes such as metabolism and membrane-generation next to the *external* processes concerned with exchanges with the environment. Our take on the pebble case seems to lack in internal counterparts to the external processes. This speaks to what the operational closure formalism indicates, which is that the pebble simply is not an operationally closed system. This means that, in terms of this formalism, there is no 'internal' to speak of that *could* operate (semi-)independently of the external processes.

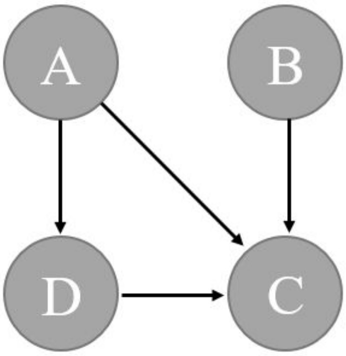

Figure 4 represents the process network relevant to a pebble on a shingle beach. The nodes with letters A, B, C and D in the figure represent the processes A, B, C, and D mentioned in-text respectively. The arrows represent enabling relations so that the arrow going down from A to C means that A enables C. Each node is coloured gray to indicate that the network is not operationally closed, as no process except for D enables and is enabled by at least one other process in the network. The network can thus not be said to be self-enabling.

5 Autonomy meets Markov blanket

Operational closure and precariousness provide the principled distinction between autonomous and non-autonomous systems. It is this distinction that seems difficult to capture within the Markov blanket formalism of the FEP: indeed, following Section 3, it seems as though both organisms and pebbles can be said to minimise free energy and can thus be cast as engaging in Bayesian inference. In Section 4 we have seen two notions from the enactive literature that are apt at capturing the difference between biotic and abiotic systems. As such, there is good reason to attempt to incorporate the enactive notion of autonomy into the FEP (Kirchhoff et al. 2018; Palacios et al. 2020).

A few FEP conceptions of autonomy exist in the literature, so it is important to discuss these and why they fall short of capturing operational closure and precariousness. According to one usage of autonomy, the internal and active states of any Markov blanketed system are considered *autonomous states*, because their values are not directly influenced by the environment (Friston, Wiese, Hobson, 2020). Yet this does not aid in a distinction between biotic and abiotic systems, as any Markov blanketed system by definition has internal and active states. One may also think the presence of active states in a system is crucial, as active states are what, in the FEP formalism, allow an organism to modulate their exchange with its environment. Yet, recall from Section 3 that a pebble also has active states. It would be strange to think that a pebble's existence has no influence on its environment merely because it does not *act* on its environment. A pebble's mass will influence the state of the water that may surround it, or the movement of the adjacent pebbles on a shingle beach, and influences the behaviour trajectories of organisms in its vicinity. These sorts of influences will be formalised as active states in the Markov blanket formalism. As such, we will be able to identify external states dependent on the pebble's active states in the same way we can do so for organisms.

Yet, one may object, a pebble's exchange with its environment is much, shall we say, simpler, than an organism's. This is roughly what is captured in the distinction between between *active* particles and *inert* particles, discussed in Friston's 2019.[8] The distinction here rests on what is called the 'information length' of a system, the technical specifics of which are outside of the scope of this paper.[9] Broadly, one could say the information length of a system corresponds to the size of the 'viable bound' of the system under scrutiny as we have discussed it in Section 2. This means that a high information length is associated with systems whose internal states display a large degree of variability, whereas low information length is associated with systems whose internal states remain largely static, or consistently revisit a very small set of states. This seems to make headway into distinguishing biotic from abiotic systems, yet fails to draw a divide *in kind,* offering only a gradual distinction *in degrees*, leaving room for a grey area between biotic and abiotic. Take the pebble, for example. For the sake of argument, let us concede that the pebble's information length is sufficiently low to be termed an *inert* particle. Yet consider now a shingle beach, consisting of a large amount of pebbles, that lies at water. We can consider the beach as a whole to have its own Markov blanket, forming an *ensemble* of the individually blanketed pebbles at the beach. The complexity of the internal states of the shingle beach as a whole as it maintains its integrity (continues being a shingle beach) in spite of the environmental fluctuations (the water flowing on and off-shore, weather circumstances, etc.) increases exponentially as we imagine it to be larger, comprising more distinct and varied pebbles, each of which 'respond' differently to the varying temperatures and kinetic forces. This increases the associated information length of the shingle beach. We could do this for increasingly complex abiotic systems until, one could imagine, the information length starts to look a lot like that of a single bacterium. The crucial point here is that relying on a system's information length may not necessarily pick out biotic systems exclusively, and remains a difference *in degree*, as opposed to a difference *in kind*.[10]

A final suggestion that could be thought to pick out organisms over pebbles is that of non-equilibrium steady-state (NESS). According to the FEP's more recent formulations[11],

---

[8] Thanks to an anonymous reviewer for pointing this out.

[9] See Friston's (2019) unpublished manuscript for a description in technical detail.

[10] What this means is that the grey area is not *inherently* a fault, yet conceding this does not help us in distinguishing the pebble clearly from the organism.

[11] Contrary to, say, Friston (2013), the condition of a system being at NESS with its environment seems to have replaced the initial clause of being locally ergodic (see Friston 2019; but also Hipólito 2020; Ramstead, Badcock et al. 2019) . Discussion of this change and its philosophical implications are outside of the scope of this paper, but see Bruineberg et al. (2020) for preliminary discussions.

any Markov blanketed system that is at NESS with its environment can be cast as minimizing free energy (Ramstead, Badcock et al. 2019). For a system to be in a non-equilibrium steady-state so defined means that the system is far from equilibrium and, *in virtue of* systematic environmental exchange, remains in the same state over time. Yet being in a steady state implies that the system remains in *the same* state over time (Gagniuc 2017). This means that, for a dynamically changing organism in constant flux, this only holds by approximation or within certain specific timeframes. An extreme example are butterflies that just got out of their cocoon, which corresponds to a massive change in the organism's states, but humans can just as well hardly be said to occupy the same state over time.[12] A pebble is in no need of environmental exchange to remain a pebble and is thus not at NESS. Yet it is also known that NESS does not uniquely pick out biotic systems (see for example Bernard and Doyon, 2015; Pourhasan, 2016). As such, it remains of import to look at the enactive approach to autonomy and how this could be approached from within the FEP.

5.1 On self-individuation

A system is considered operationally closed only if it exhibits a network of self-enabling processes. That is, each process in the network enables and is enabled by at least one other process in the network. This means that any operationally closed system is inherently composed of multiple individually distinguishable component processes. Taken together, these individually distinguishable component processes form a larger network that *self-individuates*, and generates its own boundary between itself and its environment. The Markov blanket formalism is well-equipped to capture this hierarchical boundary generation of processes (Palacios et al. 2020). If we take each component process to have a Markov blanket, and the larger, operationally closed network to have a Markov blanket too, the generation of a self-enabling and self-individuating process network can be cast as the hierarchical self-organization of a Markov blanketed *ensemble* of Markov blankets. Palacios et al. (2020) show how, with a few crucial assumptions, single cells can be shown to aggregate quite naturally into a larger ensemble. In this particular way, we can consider each node of the network to be Markov blanketed, and the ensemble-network to be Markov blanketed in itself, as shown in Figure 5 below. The nodes of the operationally closed network need not be operationally closed themselves, which means that the nodes themselves need not invite being divided further into another layered network. We can thus ground operational closure in terms of Markov blanket ensembles without inviting an infinite regress. This maps onto a single cell organism too. Consider that each organelle of a single cell can be distinguished statistically from the rest of the cell, thus establishing a Markov blanket (Palacios et al. 2020), without in itself being operationally closed and thus not in itself

---

[12] The FEP may be able to accommodate 'wandering sets' (see Birkhoff 1927) which could account for changes to a system's viable bound over time, though it remains to be seen whether this could accommodate drastic and sudden changes such as the butterfly's (Friston 2019).

requiring to be composed of a self-enabling process network under the current definition of operational closure.

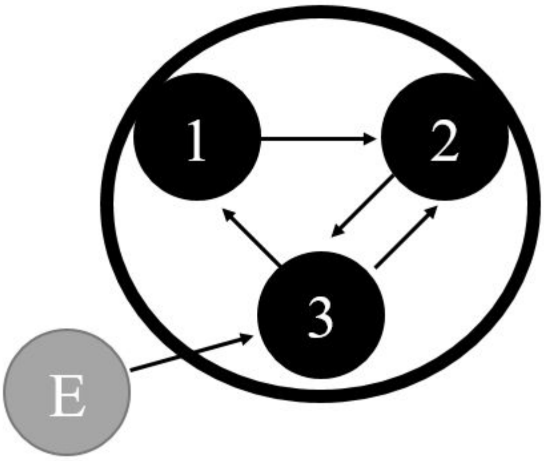

Figure 5 describes the operationally closed single cell system with a Markov blanket around the ensemble of process networks that make up the system. The process relations described are just as they were in Figure (x: cell). The circle around the self-enabling network of 1, 2 and 3 represents the Markov blanket around the ensemble.

Although this captures a key feature of operational closure, self-individuation (or membrane-generation in terms of a single biotic cell), it falls short of accounting for the conditional enabling network that differentiates autonomous from non-autonomous systems. Hierarchical self-organization is only part of the enactive story of autonomy. Indeed, the pebble challenge could be reformulated as a *shingle beach challenge* so that the beach can be cast as an ensembled Markov blanketed system that engages in Bayesian inference, composed of individually Markov blanketed pebbles. The distinction between abiotic and biotic thus remains blurred, even in a hierarchical perspective.

5.2 On operational closure and enabling relations

There are a few differences between just any Markov blanketed system and an operationally closed Markov blanketed system that we need to capture. Operational closure is a particularly *structured* manner of self-organization (Maturana and Varela, 1980). Increased structure over time implies that the long term entropy (informally a measure of disorder; Friston 2013) is low, which means that sensory surprise must be low too. The ensemble's states are constituted by the component states, which means that the component states inherit this low surprise. This is a key aspect of understanding operational closure in the FEP.

We can exploit the lower surprise internal to the network further by, for the sake of exposition, ignoring the system's environment. For each particular node, its sensory states are entirely determined by the active states of the other nodes in the network (Palacios et al. 2020). More specifically, if A enables B, that means that the active states of A must have an important influence on B, which in turn means that the active states of A are significantly determinant of the sensory states of B. Conversely, if B is enabled only by A, its sensory states are entirely determined by the active states of A. This means that, within the network, each node's sensory states are determined by the active states of its enablers. This implies that the sensory surprise of any node is at a nearly absolute minimum, *given* the active states of the enabling nodes.

In light of this, an enabling relation is closely related to the notion of *coupling*. Any two nodes can be said to be coupled when they are in a relation of *mutual* influence (Friston 2013). In active inference, the generative models associated with two coupled systems will approach one another over time, giving rise to what is known as *generalized synchrony*. As the coupled two systems continuously interact, they become attuned to one another; they adapt to one another (Friston 2013). This attunement means that the influence they have on one another becomes increasingly well accommodated. In mutual attunement, this entails changes in the *extrinsic* probability distribution in the state space, so that the sensory states associated with the active states of the coupled system are increasingly likely. On the scale of the network, this means that the nodes as part of the network, *i.e.* on a network-level scale, are in a tight coupling relation. This is to say that each node's influence will enable, and thus largely determine, another node's states that will, by virtue of being part of the network, couple back the initial node to enable and largely determine its own states either directly or indirectly. An operationally closed network, then, can be taken as a tightly coupled network of Markov blanketed nodes.

Note however, that, *prima facie*, the notion of coupling is not necessarily applicable to any two nodes in an enabling relation within the network. We thus cannot simply transcribe the enabling relations between nodes as coupling relations. A coupling relation is *symmetrical* insofar it prescribes mutual influence. This does not mean that the interaction needs to be *identical in both* directions of influence, but it does imply that the interaction is minimally bidirectional: the active states of one node determine the sensory states of another node *and vice versa*. Taken in this sense of direct influence, an enabling relation is not. An enabling relation can be asymmetrical, as we see in nodes 1 and 2 in Figure 5 above. This means that we would miss out on asymmetrical enabling relations if we were to transcribe them as coupling relations in a model. Moreover, an enabling relation concerns a specific type of influence that one node has on another. Consider that any random two systems may, for a certain duration over time, be coupled in a mutually *disruptive* fashion. This means that rather than *enabling* one another, they instead *inhibit* one another. This distinction too may be lost if we were to transcribe enabling relations as coupling relations. Crucially, however, we *can* say that the individual nodes are at least indirectly coupled to one another from a network perspective.

The network perspective can also capture the sense in which operational closure depends on a network's *constituents*. Consider, for example, how the free energy of a Markov blanketed ensemble of Markov blankets depends on the free energy of its constituents (Friston, 2013). Nonetheless, the shingle beach considerations in Section 5.1 remain valid.

5.3 On precariousness and limits

The Markov blanketed ensemble of Markov blankets has been important in our characterization of self-individuation and operational closure, so one could think it to cover precariousness as well. The idea is that an ensemble's free energy is determined by the free energy of its constituents, which means that if the constituents' free energy is minimised, the free energy of the ensemble is minimised too. It is important to note that this hierarchical dependence is a crucial feature of precariousness in organisms. That is, organisms and their component processes are inherently precarious. Yet what is typical of precariousness is not the hierarchical dependence relation: it is the natural inclination to decline.

More important here is the FEP requirement of a system to be at NESS with its environment. As we have seen, this implies that the system requires continued environmental exchange to maintain its state. *Barring the limiting remarks* of the applicability of NESS to real, living organisms noted above at the start of Section 5, the continued environmental exchange requirement for maintaining its state *is exactly what precariousness demands*. Some employ this feature of the FEP to construe cancer, for example (Manicka and Levin 2019; Kuchlin et al. 2019).[13]

Furthermore, the low sensory surprise of an enabled node, given the active states of enabling nodes may also be able to capture an organism's *precariousness*. Recall that precariousness appears on two levels in an autonomous system. Each process in the network is precarious, and the network as a unity is too. Network-level precariousness is built into the FEP at its very core. Any system needs to put in work to be able to maintain its boundaries with the environment and continue existing. This means that without this work, the system will disintegrate, which is to say the system is naturally inclined to cessation, yet remains intact due to the 'efforts' of the system. In this sense, the organism can be taken as precarious.

However, this line of thinking invites an unintended implication on the node level. Consider that high-probability sensory states are those for which they are largely determined by their enablers' active states. The cessation of a process, further, is associated with leaving expected bounds. When a process ceases, its active states will thus by definition leave expected bounds. This implies that the sensory states of an enabled node would be highly surprising (given its ceasing enabler's active states) so that it's likely to enter an unviable state and cease as well. This seems to entail that if any random enabling node would cease, the sudden

---

[13] Thanks to an anonymous reviewer for pointing this out, which inspired further considerations regarding operational closure as well.

increase in sensory surprise for the enabled nodes would sooner or later cause each other process in the network to fall like dominoes. After all, their own cessation will cause a spike in sensory surprise in the nodes they enable, and so on. As the network is composed only of processes that both enable and are enabled by at least one other node in the network, no single process will be spared. In certain cases, this is to be expected. Consider our toy description of a single cell in Section 4.1 above. If we were to cease any of the processes in that network, the entire network would collapse. Each process is essential for the continuation of the network. However, this is only a contingent fact of our toy description. As stated above, it is not necessary for each enabling process to be individually necessary or sufficient for the continuation of the enabled process. This flexibility is key in our understanding of operational closure, yet is orthogonal to the domino effect we find on a node-level of description. This shows that, though this approach is able to capture certain characteristics, it is not capable of incorporating precariousness on both a network- and a node-level of description.

Further, if we intend to capture the essential organizational dynamics for biotic systems, abstracting away the environment misses the point. By defining what something is (the system, or the *unity*), we indirectly define that which it is not (the environment) (Beer 2004; Friston 2019). This is exacerbated by the fact that for each probability distribution over internal states, there is an associated probability distribution over external states that specifies the expected influences of external states (Friston, Wiese, Hobson 2020). Even in the presence of an external environment, an operationally closed system intrinsically defines its environment as well as its boundary through which it can interact with the environment (Beer 2004 2014; Friston 2012). In an ecological situation, any one node's surprise is thus not at nearly absolute minimum, but can still be said to be *particularly* low, given the active states of its enablers.

In sum, we have presented some ways to consider conceptualizing operational closure and precariousness in terms of a tightly coupled network of Markov blankets. There is a sense in which tightly bound network-scale coupling, and particularly low sensory surprise of enabled nodes given the active states of enabling nodes, can capture operational closure and precariousness. This can be taken as a proof of concept. Further simulational research may aid further in the incorporation of autonomy into the FEP by putting the approach here to work.

Conclusion

Many FEP researchers hold the FEP to support a grand unifying ambition to account for a wide variety of phenomena, among others the organizational dynamics of living and cognitive systems. Yet a common criticism is that it is overly general and cannot distinguish between biotic and abiotic systems, making it seem uninteresting from a biological perspective. We addressed this worry by elaborating on earlier suggestions to incorporate the enactive notion of autonomy into the FEP framework. In Section 4, we described how operational closure and precariousness are concepts fit to handle the pebble challenge. In the subsequent section, we made a first attempt at incorporating the enactive language of autonomy into free energy language. We discuss different aspects of autonomy in the enactive approach and how they could potentially be transcribed into the FEP formalism. The FEP quite naturally accounts for self-individuation, a corollary of operational closure. The same applies to the bi-directional dependence relation of an operationally closed system and its component processes and a Markov ensemble and its nodes. Yet the enabling relation central to operational closure proves more challenging. There are implications with regards to the statistical relations between nodes for any operationally closed system such as an enabled node's low sensory surprise in light of its enablers' active states that we show the FEP can account for. Precariousness is also shown to be difficult to incorporate on a node-level (although the ensemble level is able to capture some basal features of precariousness), and the complexity of an ecological environment places limits on surprise-minimization descriptions as leveraged before. The FEP can thus emulate a limited version of autonomy as it appears in the enactive approach. Simulation modeling can further help incorporate this notion of autonomy into the FEP formalism.

Bibliography

Auletta, G., (2013). Information and metabolism in bacterial chemotaxis. *Entropy*. Doi: 10.3390/e15010311

Anderson, M. (2017). Of Bayes and bullets: An embodied, situated, targeting-based account of predictive processing. In T. Metzinger & W. Wiese (Eds.), *Philosophy and predictive processing (Vol. 3)*. MIND Group: Frankfurt am Main. https://doi.org/10.15502/9783958573055.

Beal, M. (2003) Variational Algorithms for Approximate Bayesian Inference. PhD thesis University of Cambridge

Beer, R. D. (2004). Autopoiesis and cognition in the game of Life. Artificial Life, 10(3), 309–326.

Beer, R. D., (2014) The Cognitive Domain of a Glider in the Game of Life. *Artificial Life, 20*, 183-206. https://doi.org/10.1162/ARTL_a_00125

Bernard, D., Doyon, B. (2015) Non-Equilibrium Steady States in Conformal Field Theory. *Ann. Henri Poincaré 16*, 113–161. https://doi.org/10.1007/s00023-014-0314-8

Birkhoff, G.D., 1927. *Dynamical systems*. American Mathematical Society, New York.

Bluck, B. J. 1967 Sedimentation of beach gravels; examples of South Wales. J. *Sedimentary Res. 37*, 128–156. doi:10.1306/74D71672-2B21-11D7-8648000102C1865D

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2018). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese*, *195*(6), 2417–2444. https://doi.org/10.1007/s11229-016-1239-1

Bruineberg, J. and Dolega, K. and Dewhurst, J. and Baltieri, M. (2020) *The Emperor's New Markov Blankets* [Preprint]. *PhilSci Archiv* http://philsci-archive.pitt.edu/id/eprint/18467

Carr, A. P. (1969) Size grading along a pebble beach: Chesil beach, England. J. *Sedimentary Petrol. 39*, 297–311. doi:10.1306/74D71C3A-2B21-11D7-8648000102C1865D

Campbell, J.O., (2016) Universal Darwinism as a process of Bayesian inference. *Front. Syst. Neurosci. 10*(49). DOI: 10.3389/fnsys.2016.00049

Calvo, P., and Friston, K. (2017). Predicting green: really radical (plant) predictive processing. Journal of the Royal Society Interface, 14(131): 20170096. Doi: 10.1098/rsif.2017.0096

Colombo, M., Wright, C. First principles in the life sciences: the free-energy principle, organicism, and mechanism. *Synthese* (2018). https://doi.org/10.1007/s11229-018-01932-w

Conant, R. C., and Ashby, W. R. (1970) Every good regulator of a system must be a model of that system, *Int. J. Systems Sci., 1*(2), pp. 89–97

Corcoran, A. W., Pezzula, G., and Hohwy, J. (2020) From Allostatic Agents to Counterfactual Cognisers: Active Inference, Biological Regulation, and The Origins of Cognition. *Biology and Philosophy*, *35*(3). https://doi.org/10.1007/s10539-020-09746-2

Domokos, Gibbons (2012) The evolution of pebble size and shape in space and time. *Proc. R. Soc. A 468*, 3059–3079. doi:10.1098/rspa.2011.0562

Daunizeau, J., den Ouden, H. E. M., Pessiglione, M., Kiebel, S. J., Stephan, K. E., Friston, K. J. (2010) Observing the Observer (I): Meta-Bayesian Models of Learning and Decision-Making. *PLoS ONE 5*(12): e15554. https://doi.org/10.1371/journal.pone.0015554

Di Paolo, E. A. (2005). Autopoiesis, Adaptivity, Teleology, Agency. Phenomenology and the Cognitive Sciences, 4(4), 429. https://doi.org/10.1007/s11097-005-9002-y

Di Paolo, E., & Thompson, E. (2014). *The enactive approach.* In L. Shapiro (Ed.), *Routledge handbooks in philosophy. The Routledge handbook of embodied cognition* (p. 68–78). Routledge/Taylor & Francis Group

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*, 129–143. https://doi.org/10.1016/j.cortex.2015.03.025

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293–301.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience, 11*(2), 127–138.

Friston, K. (2012). A free energy principle for biological systems. *Entropy*, 1392 14(11):2100–2121.

Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface*, 10(86). https://doi.org/10.1098/rsif.2013.0475

Friston, K. (2019). A free energy principle for a particular physics. Unpublished manuscript.

Friston, K., Ao, P. (2012) Free-energy, value and attractors.*Computational and mathematical methods in medicine*.

Friston, K., Kilner, J. & Harrison, L. A (2006) free energy principle for the brain. Journal of Physiology-Paris 100, 70–87, https://doi.org/10.1016/j.jphysparis.2006.10.001.

Friston, K. & Kiebel, S. (2009) Cortical circuits for perceptual inference. *Neural Networks 22*:1093–104.

Friston, K. J., Parr, T., and de Vries, B. (2017a). The graphical brain: belief propagation and active inference. *Netw. Neurosci.* 1, 381–414. doi: 10.1162/NETN_a_00018

Friston, K., and Stephan, K. E. (2007). Free energy and the brain. Synthese, 159, 417–458.

Friston, K. J., Wiese, W., Hobson, J. A., (2020) Sentience and the Origins of Consciousness: From Cartesian Duality to Markovian Monism. *Entropy 22*(5), 516; https://doi.org/10.3390/e22050516

Gagniuc, Paul A. (2017). *Markov Chains: From Theory to Implementation and Experimentation*. USA, NJ: John Wiley & Sons.

Hesp, C., Ramstead,M., Constant, A., Badcock, P., Kirchhoff, M.D., and Friston, K. (2019). A Multi-scale view of the emergent complexity of life: A free energy proposal. In M. Price et al. (eds), *Evolution, Development, and Complexity: Multiscale Models in Complex Adaptive Systems*. Springer

Hipólito, I. (2019). A simple theory of every 'thing.' *Physics of Life Reviews*, *31*, 79–85. https://doi.org/10.1016/j.plrev.2019.10.006

Hohwy, J. (2016) The self-evidencing brain. *Noûs 50*(2), 259–285. doi: 10.1111/nous.12062

Hohwy, J. (2017). How to Entrain Your Evil Demon. In T. Metzinger & W. Wiese (Eds.). Philosophy and Predictive Processing: 2. Frankfurt am Main: MIND Group. doi: 10.15502/9783958573048

Hohwy, J. (2020) New directions in predictive processing. *Mind & Language, 35* 209–223.

Kiebel S. J., Daunizeau J., Friston K. J. (2008) A Hierarchy of Time-Scales and the Brain. *PLoS Comput Biol 4*(11): e1000209. https://doi.org/10.1371/journal.pcbi.1000209

Kirchhoff, M. D. (2018). Autopoiesis, free energy, and the life-mind continuity thesis. *Synthese*, 195(6), 2519-2540.

Kirchhoff, M., Froese, T. (2017). Where There Is Life There Is Mind: In Support of a Strong Life-Mind Continuity Thesis. *Entropy, 19*(4). https://doi.org/10.3390/e19040169

Kirchhoff, M.D., Kiverstein, J. How to determine the boundaries of the mind: a Markov blanket proposal. *Synthese* (2019). https://doi.org/10.1007/s11229-019-02370-y

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, 15(138). https://doi.org/10.1098/rsif.2017.0792

Korbak, T. (2019). Computational enactivism under the free energy principle. *Synthese*. https://doi.org/10.1007/s11229-019-02243-4

Kuchling F, Friston K, Georgiev G, Levin M. (2019) Morphogenesis as Bayesian inference: A variational approach to pattern formation and control in complex biological systems. *Phys Life Rev.* doi: 10.1016/j.plrev.2019.06.001.

Kuenen, Ph. H. 1964 Experimental abrasion of pebbles. VI. Surf action. *Sedimentology 3*, 29–43. doi:10.1111/j.1365-3091.1964.tb00273.x

Laland, K., Matthews, B., and Feldman, W. (2016). An introduction to niche construction theory. *Evolutionary Ecology*, 30, 191-202.

Landon, R. E. 1930 An analysis of beach pebble abrasion and transportation. *J. Geol. 38*, 437–446. doi:10.1086/623739

Linson, A., Clark, A., Ramamoorthy, S., and Friston, K. (2018) The Active Inference Approach to Ecological Perception: General Information Dynamics for Natural and Artificial Embodied Cognition. *Frontiers in Robotics and AI*, *5* (21). doi: 10.3389/frobt.2018.00021

Manicka S, Levin M. (2019) Modeling somatic computation with non-neural bioelectric networks. *Scientific reports. 9*(1):18612. doi: 10.1038/s41598-019-54859-8

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition*. Boston: Reidel

Mossio, M. and Moreno, A., (2010) Organisational Closure in Biological Organisms. *Hist. Phil. Life Sci.*, 32 (2010), 269-288

Olivotos, S., & Economou-Eliopoulos, M. (2016). Gibbs Free Energy of Formation for Selected Platinum Group Minerals (PGM). *Geosciences*, *6*(1), 2. doi:10.3390/geosciences6010002

Parr, T., and Friston, K. (2017). Working memory, attention, and salience in active inference. *Scientific Reports*, 7: 14678 | DOI:10.1038/s41598-017-15249-0

Palacios, E. R., Razi, A., Parr, T., Kirchhoff, M., & Friston, K. (2020). On Markov blankets and hierarchical self-organisation. Journal of Theoretical Biology, 486, 110089. https://doi.org/10.1016/j.jtbi.2019.110089

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

Pourhasan, R. (2016) Non-equilibrium steady state in the hydro regime. *J. High Energ. Phys. 2016,* 5. https://doi.org/10.1007/JHEP02(2016)005

Ramstead, M. J. D., Constant, A., Badcock, P. B., & Friston, K. J. (2019). Variational ecology and the physics of sentient systems. *Physics of Life Reviews*, *31*, 188–205. https://doi.org/10.1016/j.plrev.2018.12.002

Ramstead, M., Kirchhoff, M. D., Constant, A., and Friston, K. (2019). Multiscale Integration: Beyond Internalism and Externalism. *Synthese*, 10.1007/s11229-019-02115-x.

Ruiz-Mirao, Mavelli (2007) On the way towards 'basic autonomous agents': Stochastic simulations of minimal lipid–peptide cells. *BioSystems, 91*, 374-387

Schrödinger, E., (1944) *What is life?* Cambridge: Cambridge University Press.

Skulachev, V .P., (1992) The laws of cell energetics. *Eur. J. Biochem. 208*, 203–209.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Systems Technical Journal, 27*(3):379–423.

Thompson, E. (2007). Mind in life: Biology, phenomenology, and the sciences of mind. Cambridge, MA: Harvard University Press.

van Es, T. (2019). Minimizing prediction errors in predictive processing: from inconsistency to non-representationalism. *Phenomenology and the Cognitive Sciences*. https://doi.org/10.1007/s11097-019-09649-y

van Es, T. (2020) Living models or life modeled? On the use of models in the free energy principle. *Adaptive behavior*. https://doi.org/10.1177/1059712320918678

Varela, F. (1979) Principles of Biological Autonomy. The North-Holland Series in General Systems Research, Vol. 2. Elsevier North-Holland

Varela, F., Thompson, E., Rosch, E., (1991) *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MA: MIT Press

Williams, D. (2018). Predictive coding and thought. *Synthese*, 197, 1749-1775.