3

AN ACTIVE INFERENCE APPROACH TO SEMIOTICS

A variational theory of signs

Antoine Milette-Gagnon, Samuel P. L. Veissière, Karl J. Friston, and Maxwell J. D. Ramstead

Introduction

Recent decades have borne witness to the emergence of new frameworks in psychology (Chemero, 2009; Gibson, 1979), neuroscience (Eliasmith, 2005; Varela, 1996; Varela et al., 1991), anthropology (Deacon, 2011a), and philosophy (Thompson, 2010), each more committed than the last to explaining the cognitive feats of organisms from their own point of view.

New mechanistic theories, drawing on dynamical systems theory and information theory, evince a growing interest in the study of the workings of meaning and signification. Previous theories had tended to approach living systems either from a detached, reductionist stance that disregarded conscious experience (Fodor, 1975; 1983; Watson, 1913) or – in contrast – from a purely subjective perspective, without much concern for the physical processes that underwrite our experience of a meaningful world (Ingold, 2000). The dual reality of living systems – as mechanistic, but also perspectival and capable of signification – increasingly drives research on sentient systems. It would be a mistake to think that adopting a mechanistic approach commits one to reductionism and to mindless conceptions of the living (Ramstead et al., 2018; Thompson, 2010).

The terms "signification" and "meaning" are polysemic. Sometimes their referents seem elusive because the phenomena at play in attributions of significance are multilayered. We use the terms "signification" and "meaning" to refer to a specific kind of interpretive ability and its output, whereby an agent can make use of signs to convey specific meanings – or to interpret signs provided to them as meaning this or that. As emphasized by many philosophical traditions (notably, stoicism), humans do not interact with the world in an unmediated way (as a naïve realist or phenomenological account might hold), but rather apprehend it via the meaning that they assign to their experience.

We argue that active inference can provide some mechanistic insights about signification and meaning-making abilities in humans. Active inference is the newest creature-centered framework to come out of information theory (Friston, 2020), and it represents a promising new avenue for the study of mind in cultural context (Ramstead et al., 2016; Veissière et al., 2020). Active inference explains how organisms are able to generate rolling cycles of belief-guided, adaptive action and perception (Clark, 2015; Ramstead et al., 2019a). The framework casts perception, learning, and action as forms of inference – and provides a formal model for the study of hierarchically

DOI: 10.4324/9781003051817-5

nested systems (Badcock et al., 2019a; Kirchhoff et al., 2018). The active inference framework has been applied to explain varied phenomena, such as the generation of adaptative behavior by living creatures (Friston et al., 2017), the hierarchical and nested structure of biological systems (Ramstead et al., 2019b), the construction by creatures of their ecological niche (Constant et al., 2018), and the coevolution of human genes and culture (Veissière et al., 2020), cooperative communication systems (Vasil et al., 2019), linguistic capacity (Friston et al., 2020), and hermeneutics (Friston & Frith, 2015a). We believe many of the questions addressed by active inference echo those addressed in the field of neurosemiotics, which intends to study – among other things – the processes that underwrite the general capacity of organisms to behave meaningfully, or to generate meaning, within their environment (Bouissac, 1985; Jorna, 2006).

The main claim of this chapter is that a theory of signs akin to classical, Peircean semiotics can be (partially) formulated using active inference, in light of the conceptual analogies between the two frameworks, notably the shared commitment to the notion of abductive inference in living (and cognitive) systems. Furthermore, we argue that Peirce's construct of semiotic interpretant – the interpreter's understanding of the relation between a signifier and what it signifies – can be given a formal interpretation under active inference by associating it with the construct of generative model in active inference. Generative models are probabilistic mappings, from a set of unobserved or latent causes to their observed sensory consequences. Under active inference, such models underlie the generation of adaptive action. We argue that active inference allows us to model signification and meaning-making as inference in a hierarchical generative model, where inferential processes at superordinate layers of the hierarchical model arise from – and constrain – those unfolding at subordinate levels (Friston, 2008). Based on this argument, we reinterpret Peirce's typology of signs (as icons, indices, and symbols) via active inference. We place special emphasis on the notion of abductive inference (going beyond the sensory data at hand) in resolving the ill-posed problem of inferring what our sensations connote (Seth, 2015).

The remainder of this chapter will be divided in two parts. In the next section, we introduce the active inference framework. We then detail the conceptual analogies between active inference and classical (Peircean) semiotics, before offering a partial but formal mapping of key semiotic notions, notably the icon/index/symbol triad, onto central constructs of active inference. We close with proposals for future work.

Active inference

An introduction to the active inference framework

Active inference is a theory that explains how living systems preserve their nonequilibrium steady state (Friston, 2020). According to the fluctuation theorems that generalize the second law of thermodynamics, the entropy of any system tends to increase – and living systems must counter its dispersive effects to remain alive (Parr et al., 2020; Seifert, 2012). In this context, entropy is a measure of disorder: it quantifies the total number of configurations in which the system being examined can find itself. Most inanimate systems in nature self-organize to equilibrium, which means that they consume the gradients around which they self-organize. For instance, the lightning bolt, in striking, self-organizes around – and dissolves – a charge gradient, thereby increasing the entropy and disorder of the surrounding air molecules. Living systems, in contrast, tend to resist this tendency towards entropic disorder and maintain their internal organization. By definition, this means they exist far from thermodynamic equilibrium – which is death (Schrödinger, 1944). For example, the body temperature of mammals is far from the room temperature (and for most animals, to find their body at the ambient temperature of the environment entails death). How is this feat accomplished by living organisms?

This question has been tackled by an emerging field of work centered around the free-energy principle, a variational principle of least action that underwrites active inference (Friston, 2010, 2020).

The central claim is that for a self-organizing system to exist at nonequilibrium steady state means that it looks as if it minimizes its variational free energy (Ramstead et al., 2019a). In a nutshell, this free energy quantifies the discrepancy between the current sensory state and the expected sensory state. At its simplest, this means everything that lives this must show some form of homoeostasis (Bernard, 1974) – that is, a minimization of the discrepancy between sensed states of being and homoeostatic (or allostatic) setpoints (Ramsay & Woods, 2014; Seth & Friston, 2016; Stephan et al., 2016; Sterling & Eyer, 1988). In information theoretic terms, this discrepancy is measured by free energy, which provides a measurable proxy for self-information or surprisal. The average self-information is entropy, so any minimizing free energy counters an increase in entropy (Friston, 2010). One could then reformulate the problem in the following way: "How do organisms resist entropic decay? By minimizing free energy." This is, in essence, the free-energy principle.[1] Of course, stating this without further investigation would be begging the question "How do organisms minimize their free energy?" This question can be answered by appealing to Markov blankets and generative models.

Markov blankets and generative models

We can associate a system of interest with the states in which it can find itself, which we call internal states. The rationale for this is intuitive. For a system to be considered as a system, it must evince a minimal form of independence from its surrounding environment (external states), lest it simply dissolve into that environment when perturbed – or indeed measured. This is not to say that the system of interest is completely separated from its environment (i.e., isolated or closed in a thermodynamic sense), but rather that its influence on the system is mediated, such that the embedding environment can change without the system being entrained without question.

In the active inference framework, this form of conditional independence is underwritten by the presence of a Markov blanket (Friston, 2013), a statistical construct originally developed in the context of statistical inference over random variables (Pearl, 1988). For our purposes, it operationalizes the intuitive notion of a mediated coupling to the environment (Kirchhoff et al., 2018; Ramstead et al., 2018). A Markov blanket mediates between internal and external states, and the blanket itself is composed of sensory states, which influence but are not influenced by internal states, and active states, which influence but are not influenced by external states (see Figure 3.1).

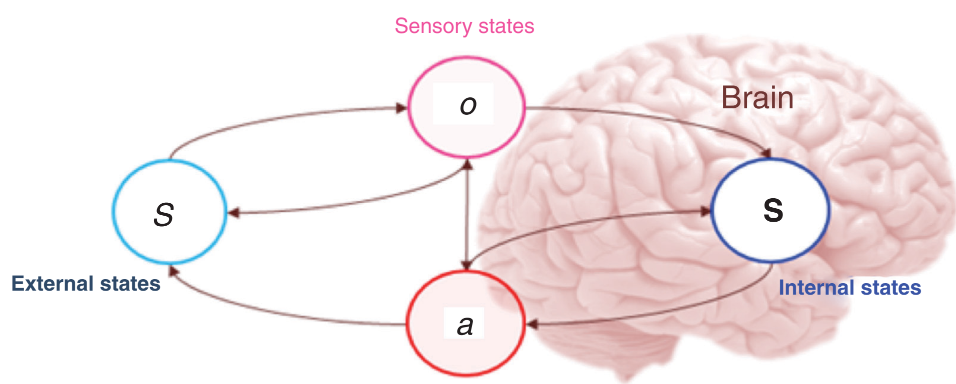

Figure 3.1 Markov blanket and its variables

This figure depicts a Markov blanket, illustrating the influences between the variables that constitute the blanket. Here, internal states – denoted s – are influenced by (but do not influence) sensory states or observations (o); while external states – denoted s – are influenced by (but do not influence) active states (a). From Ramstead, Kirchhoff, and Friston (2019). See e-book for a fullcolor version of this figure.

From this vantage point, every living creature is enshrouded by a Markov blanket, which mediates its access to the world in which it is embedded. Yet what lies behind the shroud is of critical significance to the creature because everything of relevance to it – be it potential predators, food items, or mates – lies there. This is reminiscent of the concept of the *Umwelt* developed by Jakob von Uexküll. The *Umwelt* can be understood as the embodied perspective (i.e., the viewpoint) that biological beings take on their world, and this perspective is the union of the *effector world* and the *perceptor world* (von Uexküll, 1982). This means the subjective world of organisms is not composed of a series of objects with "objective" properties, but rather is better understood as a set of objects that are made meaningful in the organism's perceptual field by its action capacities: for foxes, given their specific set of hunting capacities, a field of grass is a support for walking towards their prey; but for worm-eating birds, given their action capacities, the field grass is the "pantry" itself, so to speak.

The existence of a Markov blanket implies a form of inference, implicit in the minimization of free energy that scores the likelihood of sensations, given an *Umwelt* associated to the internal states. This means that to exist – in the sense of being separable from the environment – is to *infer* the external causes of sensory impressions that are hidden behind the Markov blanket. These hidden causes are the external states. In active inference, inference about the most probable hidden cause of sensory data is accomplished thanks to what is known as a generative model, a statistical formulation of the *Umwelt* as a model of the process that generates sensations. In other words, it is a model of how sensory impressions are generated – including, crucially, the actions of the organism. Technically, a generative model is a joint probability distribution or density over sensory outcomes and their hidden causes (Friston, 2010; Friston et al., 2018).

The idea is that organisms entertain and evaluate competing models or hypotheses (*Umwelten*) and select the one for which there is the most evidence (i.e., the hypothesis that renders the sensations the most likely or least surprising). Crucially, this evidence is the complement of the free energy above. In other words, minimizing free energy minimizes the discrepancy between sensory states and those expected under a generative model. This is mathematically the same as maximizing the evidence for the generative model. Heuristically, variational free energy therefore provides a measure of the evidence for a creature's generative model. To select the model that minimizes the discrepancy between what is predicted and what is sensed (i.e., variational free energy) is the same as selecting the model that is supported by the most evidence (Friston, 2020) – that is, a form of self-evidencing (Hohwy, 2016). The links between this self-evidencing and Peircean abductive inference will be fleshed out below.

We emphasize the inferential and perspectival nature of active inference. Generative models scaffold the probabilistic beliefs of an organism. These Bayesian beliefs[2] reflect the existence of the organism itself. They have been shaped by evolutionary history: the preferred states of the organism, encoded in its model, can be cast as a "best guess" about the causal structure of the econiche as experienced by the phenotype in question (Badcock et al., 2019b; Campbell, 2016). In other words, organisms are constantly trying to infer states of affairs in the external world, given current evidence provided by sensory states and its prior beliefs – and, indeed, given the kind of creature that it is. For instance, for a human to find oneself suddenly submerged in water would be quite surprising, whereas such a situation would be expected for a fish. What is important to note is that the generative model is not an objective map of the world; it is formed as a function of the needs and concerns of the creature (Ramstead et al., 2019a; Tschantz et al., 2019). Thus, generative models harness these species-specific expectations about the lived world (just like in the *Umwelt*).

A simple generative model: Likelihood mappings and prior beliefs

Given sensory data, the system is trying to infer what external states have caused these outcomes. The relation between hidden causes and sensory outcomes is formalized in a generative model through a likelihood mapping (Figure 3.2), which in discrete-state space generative models, like

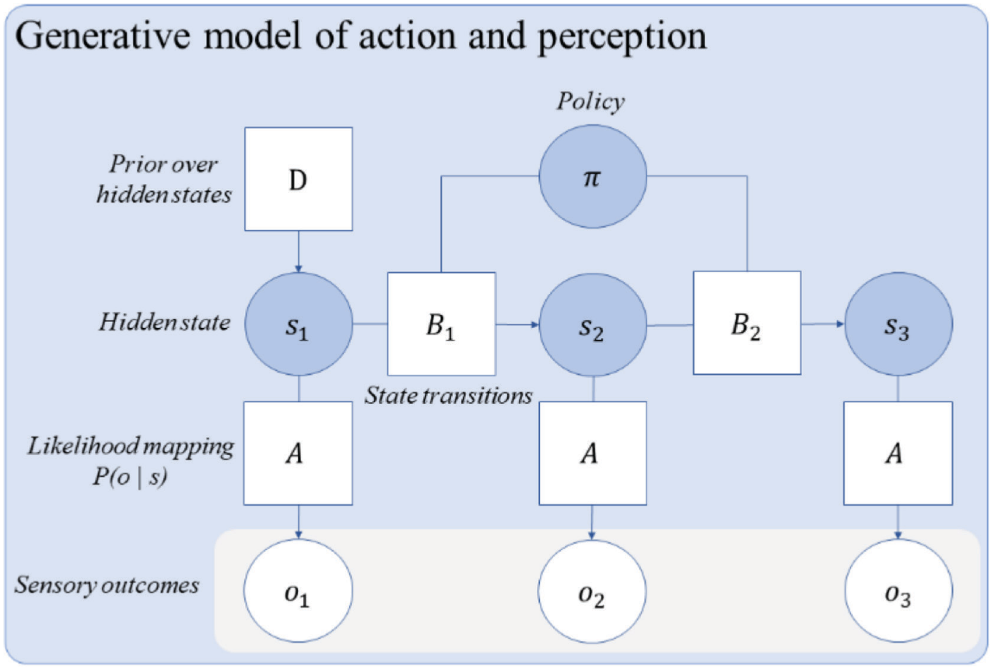

Figure 3.2 Generative model of action perception

In active inference, perception amounts to the (Bayesian) inversion of a generative model. This inversion corresponds to inference or abduction. The organism has access to its sensory observations (o), its beliefs about how its observations map onto states in the world (A), and its prior beliefs about the state of the world before seeing observations (D). More sophisticated generative models include beliefs about state transitions (B) that depend upon plans or policies. The policies (π) that are selected are those that minimize expected free energy that depends upon prior beliefs about final states (C). Please see Parr and Friston (2018) for a technical explanation of the equations that describe the implicit belief updating, given sensory outcomes. See e-book for a full-color version of this figure.

finite state machines or partially observed Markov decision processes is usually denoted by a matrix A. Technically, this mapping tells us the probability of some observation, given that some state of the world is the case – which is denoted P (o | s). This likelihood is supplemented with some prior Bayesian beliefs about how states change over time, usually encoded by a probability transition matrix called B. These state transitions depend upon action or plans. Prior beliefs about final and initial hidden states are usually encoded in vectors called C and D. It may seem strange to reduce a generative model, or *Umwelt*, of the world in this way; however, the functional form of this generative model is universal and very expressive (see below). Crucially, committing to a particular functional form allows one to simulate the minimization of free energy and accompanying self-organization. In this setting, the minimization of free energy corresponds to Bayesian belief updating and, with the above form, looks very much like message passing in neuronal circuits (for further details, see Parr and Friston, 2018).

Starting from prior beliefs and its sensory states, the organism must solve the inverse inference problem that of determining which state of the world is most probable, given its observations and prior beliefs – denoted P (s | o). Thus, we say that perception "inverts" the generative model that maps from unobservable (external) causes to observable (sensory) consequences to obtain the inverse mapping – namely, the most probable cause of current sensory states. This inverse mapping from consequences to causes is the essence of inference and self-evidencing, which can also be viewed as an act of abduction.

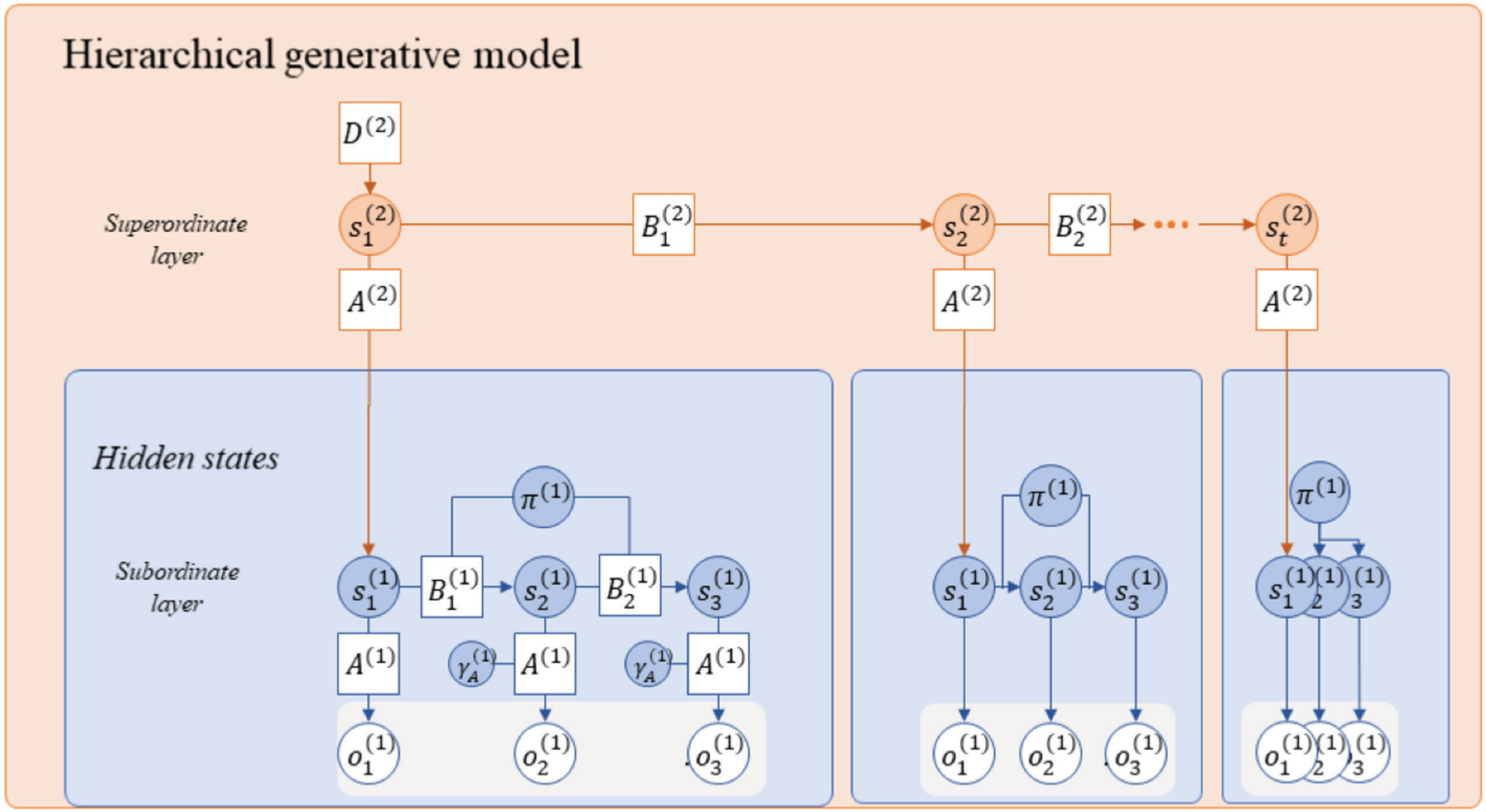

Figure 3.3 Hierarchical generative model

The generative models presented here can be hierarchically extended. In such hierarchical models, state estimation at the subordinate level is used as sensory evidence (i.e., observations) for state inference at the superordinate level. Please see Parr and Friston (2018) for a discussion of the implicit belief updating dynamics. See e-book for a full-color version of this figure.

Crucially, to be endowed with beliefs about state transitions opens the door to action on the world. After all, the ability to act depends on the ability to anticipate the (sensory) consequences of action, which in turn implies the ability to plan into the future. Under active inference, beliefs about possible courses of action are called policies. The selection of policies is implemented as the selection of specific beliefs about state transitions – namely, beliefs informed by the foreseen consequences of action and prior beliefs about final outcomes. Action is cast as a self-fulfilling prophecy: creatures fulfill their expectations through policy selection (Friston et al., 2017). Policy selection, in turn, is driven by expected free energy – namely, the free energy expected under a given policy. Differences in expected free energy over competing policies drive policy selection, as creatures enact the policy associated with the lowest expected free energy, thereby working towards predictable outcomes that constitute the free energy minimizing, evidence maximizing, nonequilibrium steady-state dynamics that characterize the creature in question.

Crucially, for our purposes, the generative models used in active inference can be endowed with a hierarchical structure. Such a scheme is depicted in Figure 3.3, in which state estimation at the subordinate level is used as sensory data (as observations) for state inference at the superordinate level. Note the higher-order likelihood matrix A(2), which links lower-level states (1) – now taken as sensed outcomes – to higher-level state inference (2). This setup effectively equips the agent with higher-order beliefs, e.g., higher-order beliefs about state transitions, which would be denoted by B(2), that contextualize beliefs and inferences at the subordinate layers of the model. In short, a hierarchical generative model of this sort equips the creature with a sense of the future (i.e., temporal depth) at different temporal scales, which finesses planning and policy selection – based upon the future consequences of action.

Toward a variational semiotics

We are now in a position to establish a preliminary mapping between semiotic constructs and structures within the active inference framework. We first expand the general theme of our variational rendition of semiotics by showcasing the conceptual similarities between the two frameworks. More precisely, we claim that Peircean interpretants can be associated formally with generative models in active inference. We then describe how Peirce's most famous typology of signs into icons, indices, and symbols can be recast in the active inference framework.

As foreshadowed in our brief presentation of active inference, we note that the theoretical structure of active inference is strikingly similar to that of Peircean semiotics, allowing us to map some of the constructs of active inference (at least partially) onto those of Peirce's semiotics, namely abductive inference and the sign triad of icons, indices, and symbols. Does that mean we can *reduce* semiotics to active inference? Our aim is rather to show that active inference is semiotic in its logic. In some sense, our attempt to explain some of the semiotic concepts by appealing to the resources of the active inference framework is not far from the intent of Peirce himself. As Short argues, "a central thrust of Peirce's mature semeiotic is that intentionality may be explained naturalistically" (Short, 2007: 8).

Active inference and deflationary semantics

Before touching on the semiotic structure of active inference, we have to address one major concern about generative models, which is their capacity to address the very notion of "meaning" that is at stake within neurosemiotics. Indeed, some critics have pointed out that information-theoretic constructs should not be conflated with the phenomenon of "meaning" (Brier, 2008). Thus, it is fair to ask whether active inference and generative models are up to the task. Our goal here is not to demonstrate that a global reduction of the phenomenon of meaning is within the purview of active inference, but rather to point toward how active inference can deal with certain specific aspects of meaning. As stated before, the term "meaning" itself is polysemic; one could follow Morris and distinguish between multiple dimensions of meaning, namely *semantics* (or the "reference" or "aboutness" of signs), *syntactics* (the reference to other signs or symbols) and *pragmatics* (the concrete use of signs in real-life situations).[3] This leaves out the *phenomenal* or *experienced* aspect of meaning, arguably infusing all other dimensions to different degrees. While we do not pretend to directly solve the issue of *naturalizing* phenomenality as the complexity of the issue as generated more than 100 years of active philosophical and scientific debate (see Ramstead, 2015 for a discussion), this chapter still aims to convince the reader that active inference can at least provide some partial account of the "meaningful" drive of organisms that can be also described as dynamical systems.

The *semantic* aspect of meaning within active inference has been studied by Ramstead et al. (2020). In the active inference conceptual context, *semantics* is used to describe the relation between internal states of the system and external states – that is, an *aboutness* of internal states (and not meaning as a subjective feeling or experience). The authors, following Egan (2019), propose to cast, in a deflationary fashion, the aboutness of internal states in terms of the mathematical function that these states accomplish within physical systems, rather than by some "gloss" of cognitive content. The key to understanding how this (mathematical) aboutness is at play is that we can describe a given system with the tools of systems theory, as a dynamical flow within over some state space, or using information theoretic tools, providing a statistical description of this flow. In brief, under the free-energy principle, we can describe the internal states of a system as containing a *statistical manifold*, a space in which the coordinates are *sufficient statistics* of beliefs over external states (e.g., the mean and variance of a Bayesian belief).

Under the free energy principle, given the Markovian partition of internal and external states (which are conditionally independent), it follows that internal states act as (the *sufficient statistics* of) beliefs about external states – that is, given the internal state, a consumer of the state can find the probability distribution of external states (just like any normal distribution can be reproduced given its mean and variance, which are its *sufficient statistics*). In short, the *aboutness* of internal states is formalized under the free-energy principle as a relationship between internal and external states, with internal states *parametrizing* probability densities over external states – a technical treatment of the mathematical foundations of this observation can be found in Friston (2020) and Friston et al. (2020). Philosophically, the important point to understand is that, in this context, the emphasis is placed on the *relations of meaning* rather than on the *content of meaning.* Furthermore, in active inference, the relations of meaning are always enacted: the internal states also encode the *consequences* of action on external states (via active states).

The limitation of this account is that we have to accept a "deflationary" account of semantic content (Egan, 2019), in which semantic content is only an intentional gloss that can serve heuristic purposes to guide scientific research about cognitive activities (Ramstead, 2019c). That is, the association of internal states with external states and the role of this association for the description of the behavior of a far-from-equilibrium steady states system can be explained mathematically, but not the complete integrated experience of meaning. Nevertheless, we still maintain that active inference displays some basic aspects of a semiotic structure. With these deflationary caveats in place, we now attempt to connect systematically active inference and Peirce's semiotics.

A basic semiotic structure

Signs are first and foremost things that stand for other things. This is in line with the commitments of active inference, which states that creatures do not have access to things in an unmediated way (as we have presented in comparison with Uexküll's *Umwelt*). This is why we believe that, at its core, active inference is compatible with many aspects of (Peircean) semiotics. Indeed, the abductive nature of the inferential processes under the free-energy principle has recently been compared with Peircean abduction (or abductive inference) (Pietarinen & Beni, 2021). Briefly, they describe self-evidencing processes in the following fashion:

1. Y (a sensory datum) is surprising (in the information-theoretical sense described above);
2. if X (a prior or a fact about the world) were to be the case, then Y would be less surprising;
3. therefore, let us see to it that/there is reason to suspect that X (to "believe-X" about external states) is to be part of the generative model. (Adapted from Pietarinen & Beni, 2021; see also Parr and Friston, 2018, for a technical explanation of the equations that describe the implicit belief updating, given sensory outcomes.)

The inferential nature of semiosis has also been described within the field of biosemiotics. For instance, Kull et al. (2011) argue that conditional forms of logic (i.e., if → then) are embodied in the forms and habits of organisms and are in this sense a "bio-logic." As they explicitly state,

> semiosis facilitates the development of an organism's capacity to behave in a way that is both consistent with its environment and implicitly inferential. "Logic" as we are using it here is not something to be considered as a product of abstract cognition in humans, but rather we simply intend to highlight the inference-like architecture of biological function, which we take to also be the basis of semiosis in general.
>
> *(Kull et al., 2011: 33)*

This "inference-like" architecture is precisely what is at play within active inference under the free-energy principle, following the fact that internal states of organisms are conditionally independent of external states, given the Markov blanket partition.

In this line of thought, we propose that the Peircean concept of *interpretant* can be mapped to the concept of generative models, an idea that has already been sketched, but not developed enough, in Campbell (2012). The interpretant, for Peirce, is the "understanding" (at least implicit) that a semiotic agent has of the signifier–signified relationship – or, as Peirce puts it, the effect of the sign on the interpreter (Peirce, 1902). This construct reflects the idea that a signifier must be taken as standing for something. That an agent can relate a signified to its signifier assumes that they have acquired the relevant skill, or evolved the disposition, to interpret the sign as a sign. At bottom, a Peircean interpretant is a way to sample, and organize our perception of, the world around us.

Generative models accomplish the same thing as interpretants – namely, organizing perception of (and action in) the world. Indeed, generative models are joint probability distributions of sensory states and what caused them, including the action of the organism itself. In other words, the inferential process at the core of life – the *bio-logic* as described by Kull and colleagues – takes the form of a generative model within active inference. This affinity between semiotic and variational frameworks makes us believe in the possibility of a (partial) formalization of semiotics as a *naturalized* theory of signification. The aim here is not to replace the field of semiotics, but rather to suggest that semiotic processes can be explained rigorously by appeal to active inference framework. Moreover, this correspondence can facilitate the work of understanding semiotic constructs at different temporal scales because active inference provides a principled (and possibly unified) account of life processes, including signification and semiosis.

Variational icons ("looks like")

We now turn to a partial breakdown of the Peircean typology in variational terms, starting with the simplest signs: icons. Icons signify by virtue of a shared quality (often an apparent physical similarity) between the sign-vehicle (signifier) and the object (signified) (Peirce, 1902). An illustrative example of an iconic sign is a photograph of a friend. Unless affected by prosopagnosia or placed under undue strain, persons can leverage the palpable similarity between the photograph and their friend and recognize the image as being a photograph *of* this friend.

Iconic signs underlie the most elementary forms of perception, since members of a given perceptual category bear iconic resemblance to each other, for the interpreter. Perceptual categorization is thus dependent on iconicity – in the general sense of shared quality or similarity – because a category is simply a set of members sharing certain features, which means that they signify each other iconically (for the interpreter who uses the category).

From the variational point of view, we argue that a sign signifies iconically when its action on the interpreter is mediated by a likelihood mapping (or matrix A). Thus, two different icons refer to the same thing (object, event, etc.) if they trigger the same inferences. If a given signifier is similar enough to what it signifies, it will cause the creature to infer the same hidden cause as when it observes the thing signified by the icon. In short, iconic signifiers trigger the same state inferences as the direct, firsthand perception of their signified "in the flesh": both the signifier and the signified itself map onto the same underlying latent state. Our proposal here is similar to Deacon's description of icons: the most basic sense of iconicity is that of "non-distinction" – that is, that the same interpretative processes are in play for two things, making them "iconic" of each other (Deacon, 1997: 76–77). The neuroanatomical correlates of this matrix A function will be explored in the next section, linking icons (A matrices) and indices (B matrices).

Variational indices ("points to")

Indices can also be formalized via active inference, through the correspondence between the process of indexical inference and beliefs about state transitions (i.e., B matrices). The capacity of indices to signify is based upon causal or correlative relationships between them and their signified. For example, because of the causal relation between them, symptoms can be indices of a disease. Here, we follow Deacon's hierarchical model of semiotics, in the sense that simpler types of signification are embedded within more complex types – that is, there is a nested structure of signs at play in the constitution of more complex semiotic relations. To recognize a sign as the index of some signified, the creature needs to recognize (iconically) the two elements composing the indexical relation (e.g., symptom and disease) and relate them using a superordinate sign. We argue that this structural (and hierarchical) characteristic is similar to the way that, in active inference, moment-to-moment state inference (based on likelihood mappings and prior beliefs) is contextualized by beliefs about state transitions.

Recall that state transition (B) matrices embody – in neuronal or chemical connections such as the brain connections or a cell's intracellular kinetic pathways – a creature's beliefs about how states transition into others, and about what sensory outcomes these future hidden states typically cause. Moreover, when one is talking about a causal relationship between things (one of the hallmarks of indexical signification), one is really talking about transitions between one state and the next, at least insofar as this relation of co-occurrence is believed to hold by the interpreter. What we mean is that causality is not necessary "objective" – for example, a lab mouse can have a model of the transition between a red light and an electric shock, whether or not the red light really *causes* the electric shock. Therefore, we argue that the transition probabilities (the B matrices) function, in semiotic terms, indexically, since they imply a succession of moment-to-moment state inferences based on likelihood mappings and priors, just as indices imply an association of icons. This is particularly so for generative models based upon discrete state spaces, where the only thing that distinguishes one state from another is its index – and the only operational meaning of these indices are the transition (or likelihood) mappings to the indices of other states (or sensory outcomes) – see Friston and Buzsaki (2016) for treatment of indexing time in this setting.

Interestingly, the general structure of these (perceptual) inferences can be related to neuroanatomical architecture (Hipólito et al., 2021; Parr & Friston, 2018). This is in line with the good regulator theorem (Conant & Ashby, 1970), stating that a good regulator must be a model of what it regulates. For active inference, this means that generative models and brain anatomy are mutually constraining – that is, that the space of plausible brain architectures is constrained by the free energy principle (Parr & Friston, 2018).

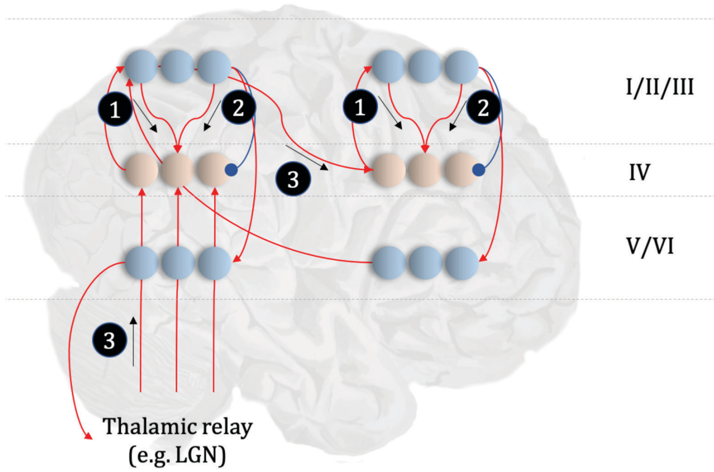

Figure 3.4 Active inference model mapped onto a brain diagram

Neurons in cortical layer IV represent the spiny stellate cells that receive input from relay nuclei of the thalamus, and from lower cortical areas. The appropriate thalamic relay depends upon the system in question. In the context of the visual system, it is the lateral geniculate nucleus (LGN). In the somatosensory or auditory systems, it is the ventral posterior nucleus or the medial geniculate nucleus, respectively. Layer IV cells in this network signal prediction errors (a measure of surprisal), computed by comparing the optimal estimate (obtained by combining the messages from its Markov blanket) with the current belief, represented in superficial cortical layers (adapted from Parr and Friston, 2018). See e-book for a full-color version of this figure.

Active inference models have been mapped onto known neuroanatomy. Consider the simplified model of the brain presented in Figure 3.4, adapted from Parr and Friston (2018). In this representation, the superficial layers of the brain encode an agent's expectations: the units in superficial layers encode beliefs about states at three time-points (immediate past, current time, and immediate future). In this neuronal representation, the likelihood mapping is encoded by specific neural populations (which implement the A-matrix that as we have seen encodes the probability of an observation given a discrete hidden state) and passed on via the thalamic relay (by the lateral geniculate nucleus for the visual system or the medial geniculate nucleus for the auditory system). Conversely, the neural messages carrying signals encoded by B matrices (state transitions) are passed on and processed by populations encoding past states and future states, which are thought to be implemented by the connections between superficial layers and layer IV spiny stellate cells. In turn, these neurons are thought to represent prediction errors, which as they arise and are quashed drive the system towards better predictions. This forms a biologically plausible neural architecture that can implement icons and indexes as discussed above.

Let us use an example to see the possible implications of such framework. Imagine a creature walking in a forest – this example is inspired by Deacon (1997: 77–78). Suddenly, a new sensory outcome is made available – one that updates its current belief about the state of affairs in the world. After correcting for prediction error, the creature infers (that is, *believes that*) what it is seeing is smoke. This first (predictive) inference allows the agent to attribute the changes in its sensory states to a cause – namely, to a black column rising from the trees. Armed with the likelihood mappings already formed within the organism's generative model, the animal infers that smoke is the most probable hidden cause of the current sensory outcomes. Here, as Deacon suggests, smoke is perceived iconically, in the sense that the current appearance of the black column bears an iconic resemblance to past occurrences of smoke. As we saw in the previous section, active inference formalizes this kind of relationship with likelihood mappings. The organism can make this inference because it has a good generative model of an econiche that features things that "smoke."

With the presence of the smoke inferred, other beliefs get to work. Through its previous experiences with its niche (made of combustible trees), the organism has become endowed with beliefs about state transitions (B matrices). Now that smoke has been identified as the probable cause of the current sensory outcome, uncertainty can be reduced further by leveraging these beliefs. Our creature has learned that *the most probable cause* of smoke is fire. It is thus able to perform a nested indexical inference (namely, that the state "smoke" often transitions into the state "fire"), so to perceive smoke motivates the inference that there is probably a fire nearby, with (action) related sensory outcomes associated with the hidden state "fire" (e.g., heat, light, burning, pain), thus motivating policy selection to ensure movement away from the source of smoke (even if there is not a *real* fire).

Variational symbols ("refers to, within a convention")

We now turn to symbols, the most complex case in the semiotic theory of signs. A symbol signifies by virtue of a convention (Peirce, 1902). Although this definition is generally agreed upon, it leaves much to the imagination. This definition is vague enough to lead to multiple usages of the term. Thankfully, the construct was taken up by Deacon (2011b), who characterized two major usages of the word "symbol" in the literature:

1. *Non-linguistic symbols.* In the social sciences, the word "symbol" refers to meanings that are conventionally invested in or projected onto artifacts that are culturally determined. The meaning is typically learned and culturally specific. For example, crowns are symbols of monarchy and its related political institutions, as when we refer to a monarchy as "The Crown."
2. *Linguistic, code-like symbols.* In mathematics and logic, the word "symbol" refers to written traces that are conventionally mapped to other such traces, and that can be combined to form other, distinct written traces according to explicit rules.

We argue that (1) the conventional aspect of symbols can be understood as an effect of collective patterns of inference based on the same, shared generative model; and (2) the syntactical aspect of symbols (symbol to symbol transitions, as are commonplace in language) can be recast as superordinate beliefs about (semiotic) state transitions.

Deontic cues and the conventionalization of signs

We believe that studying the *conventional* aspect of symbols is akin to studying cultural patterns inside a given group or society – that is, studying a multi-agent phenomenon. How can the active inference framework explain these phenomena? The trick is to reformulate the question: How can multiple agents arrive at a shared understanding of the meaning of some cues in their (shared) environment? We believe we can recast the conventional aspect of symbols as *deontic cues under a shared generative model*.

Indeed, what we call "culture" – that is, the inheritable behavioral differences among conspecifics that are acquired through learning – can be understood as underwritten by shared generative models. Thus, to share a culture is to share a set of expectations about "how creatures like myself behave in specific contexts," a model of the "generic other" reminiscent of Mead (1934). This notion of a shared generative model, in turn, has motivated work on shared manners of attending to the world – what we have called regimes of attention and expectation (Constant et al., 2019). These comprise shared manners of sampling of the world and shared patterns of attributing salience to things in the world.

In a very minimalist sense, a cultural practice can be described as a shared, socially patterned way of acting and perceiving the world (Veissière et al., 2020). In other words, people sharing the same culture will share the same (or similar enough) generative models, which means they share the same expectations about how the world is and how agents can act within it. Shared regimes of attention manifest themselves to the organisms through shared behavioral repertoires underwritten by shared expectations about the value of policies; these are *cultural* affordances because they are not innate, but instead learned through immersive practice and imitation.

How can a generative model become "shared" among multiple agents? For simple agents, like songbirds, the *attunement* of two (or more) generative models can be described as "neural hermeneutics," which describes the process by which two interpreters can come to understand one another – a process that rests on sharing a generative model. In short, two coupled interlocutors attune themselves to each other until they converge on the same beliefs about each other's internal states (Friston & Frith, 2015a, 2015b; see also Veissière et al., 2020). For humans (and any other potential *cultural* species), the story is, of course, more complicated.

A key notion when it comes to understanding these phenomena is that of *niche construction under the free energy principle* – that is, the modification of the environment by the organism to better fit its own expectations, as formalized within the active inference framework (Constant et al., 2018). Indeed, active inference is not a process situated solely in the brain, but rather is embodied and embedded in the environment. The modifications of the environment by the organism are targeted towards elements that can be modified to better fit its own phenotype, and thus minimize its free energy in that environment. These modifications are "traces" of the action of the organism; the environment embodies the preferences of the organisms that shape it. Reciprocally, that means the organism can later rely on environmental features to guide its behavior (or policy selection), a concept known in the philosophy of biology as the *scaffolding* of cognitive processes (Sterelny, 2010). Under the active inference framework, policy selection – involving environmental features – can have an *epistemic value* or a *pragmatic value* (Constant et al., 2019). *Epistemic value* is a function of the reduction of uncertainty, while *pragmatic value* is a function of the potential fulfillment of preferred sensory outcomes (innate or prior preferences over phenotypic states of being).

The more organisms that act on its environment in a certain way, the more the environment becomes *robust* to change (like a path in a field whose groves grow deeper and cleaner at each passage, and thus comes to "embody" the preferred path of the individuals who take it). In this context, generic "cues" (i.e., iconic/indexical signs) that have an epistemic and/or pragmatic value (e.g., like the link between smoke and fire) will consolidate into *deontic cues* (i.e., conventional signs) as a function of the action of the organisms within the environment (Constant et al., 2019). A *deontic* cue informs the organism about what ought to be done given the current context and situation, and is socially determined. This is because the more organisms act on their environment, the more their actions are carved into their niche – and the more the niche will reliably inform other members about the expectations of a *generic like-me other* in this (now socio-cultural) niche. Hence the opportunity for deontic policy selection, where denizens scaffold their inference through the environment in a reciprocal causation loop. Thus, the environmental niche itself becomes a model of its inhabitants, in the sense that this reciprocal loop performs a "caching" of beliefs about actions in this environment. In other words, *generic cues consolidate into deontic cues through agential actions on the environment*.

Conventions – that is, shared ways of acting and perceiving the world – are thus generated through the actions and the recognition of the action of the others, eventually amounting to a local "social-cultural world" where mutually recognized utterances and environmental cues have a specific "meaning," or a certain epistemic and pragmatic affordance (or value) (Veissière et al., 2020). This is also reminiscent of the Morris *pragmatics* dimension of meaning as described earlier. The deontic cues that make up the niche encode the preferences of the "generic other" and can take the form of indexical artifacts that acquire a deontic value through the repeated actions of multiple agents and *point to* broader symbolic conventions for navigating the world. Roads, for example, point to the expected presence of other humans – who also know how to operate cars – to be utilized in certain explicit ways (via traffic signs as additional guides) and many more implicit ways encoded in cultural conventions. We believe that these deontic cues – as characterized within active inference – capture the conventional aspect of symbols – that is, environmental and cultural cues that have a shared value and will influence and be interpreted by different members of a community in roughly the same way. Technically, in Peircean nomenclature, this conventional feature is characteristic of *legisigns* – that is, sign-vehicles that are generated and used conventionally, whereas symbols are a specific kind of legisigns (see Short, 2007 for a more detailed account of the Peircean typology of signs.

The syntactic aspect of symbols

The second aspect of *symbols* is more akin to the linguistic notion – that is, signs that can indicate other signs or, in the hierarchical model of Deacon, *indices of indices* (Deacon, 1997). Based on generative models with deep temporal structure (e.g. Friston et al., 2020), we suggest that the best way to implement syntactic constraints in a generative model is to use a hierarchical model that has a superordinate B matrix. It has already been argued that statistical notions could yield a working paradigm for understanding language processing, acquisition and evolution, which could provide an alternative to the innateness of language structures (Christiansen & Chater, 2016). Indeed, what is a syntax if not a set of rules ascribing the correct and incorrect transition from one symbol to the next, and which symbols can indicate which other symbols. An important feature of (linguistic) symbols is the fact that they can refer to other symbols. This can be described as a syntax (Deacon, 1997; Luuk & Luuk, 2012). Interestingly, it is precisely this kind of hierarchical and relational indexing that underlies most of modern computational linguistics (Khani et al., 2018; Kleinschmidt & Jaeger, 2015; MacKay & Peto, 1995; Roy, 2005; Teh et al., 2006), sometimes with an explicit nod to semiotics.

With a generative model of two systems able to play the 20 questions game, Friston et al. (2020) suggest that transition between words in a sentence is analogous to policy selection – namely, selecting the best symbol given precedent symbols, future symbols, and the general context of the conversation. Symbols can be recast as state transitions, beliefs about state transitions beliefs, or B matrices at a superordinate level of the generative model, which contextualize subordinate B matrices. This is the variational version of indices of indices. This refers to transitions between symbols (e.g., selecting which word will come next in a sentence). This rendition is, of course, not a finished model of all linguistic processes that, in itself, could count as a dynamical explanation of syntax, but rather functions to illustrate the potential of active inference to construct such a model.

Conclusion

We hope we have demonstrated the possible fruitful connections that can be made between the field of semiotics, which provides tools to understand signification, and the active inference framework, which provides tools to understand living systems from a formalized and principled statistical perspective. The work presented here is a first sketch of what could be called variational semiotics. We hope this discussion will motivate interdisciplinary investigations at the intersection of semiotics and the sciences of life and mind, progressing towards the establishment of neurosemiotics to understand the unyielding drive to construe meaning in humans (and possibly other species).

Acknowledgments

We are grateful to Casper Hesp, Pierre Poirier, and Simon Tardif for helpful comments and discussions that contributed to our work on this chapter. We are also grateful to Lars Sandved Smith for permission to use his figure template for Figures 3.2 and 3.3 and Thomas Parr for permission to use Figure 3.4. Researchers on this article were supported by the Fond de recherche du Québec – Société et Culture (AMG), Social Sciences and Humanities Research Council of Canada (MJDR) and by a Wellcome Trust Principal Research Fellowship (Ref: 088130/Z/09/Z) (KJF).

Notes

1 Technically, the free energy principle turns this on its head to show that systems that resist entropic decay (i.e., possess a nonequilibrium steady state) look as if they are minimizing free energy, thereby furnishing a teleological normative account of self-organization.

2 Probabilistic or Bayesian beliefs are simply conditional or posterior probability distributions over external states that are parameterized by internal states. They do not connote personal or propositional beliefs. For example, a virus can encode Bayesian beliefs about its external milieu in its internal molecular states.

3 This characterization of Morris has been described as non-Peircean (Pietarinen, 2012); nevertheless, it proves useful in our situation to relate certain aspects of meaning with certain aspects of active inference.

References

Badcock, P. B., Friston, K. J., & Ramstead, M. J. D. (2019a). The hierarchically mechanistic mind: A free-energy formulation of the human psyche. *Physics of Life Reviews*, *31*, 104–121. doi: 10.1016/j.plrev.2018.10.002

Badcock, P. B., Friston, K. J., Ramstead, M. J., Ploeger, A., & Hohwy, J. (2019b). The hierarchically mechanistic mind: An evolutionary systems theory of the brain, mind and behavior. *Cognitive, Affective, and Behavioral Neuroscience*, *19*, 1319–1351. doi: 10.3758/s13415-019-00721-3

Bernard, C. (1974). *Lectures on the Phenomena Common to Animals and Plants* (H. Hoff, R. Guillemin, & L. Guillemin, Trans.). Springfield, IL: Charles C. Thomas.

Bouissac, P. (1985). Neurosemiotics: A definition. *Recherches sémiotiques/Semiotics Inquiry*, *5*(3), 323–325.

Brier, S. (2008). *Cybersemiotics: Why Information is Not Enough!* Toronto: University of Toronto Press.

Campbell, J. O. (2012). Bayesian inference and the world mind. https://arxiv.org/pdf/1210.8031.pdf

Campbell, J. O. (2016). Universal Darwinism as a process of Bayesian inference. *Frontiers in Systems Neuroscience*, *10*. doi: 10.3389/fnsys.2016.00049

Chemero, A. (2009). *Radical Embodied Cognition*. Cambridge, MA: MIT Press.

Christiansen, M. H. & Chater, N. (2016). *Creating Language: Integrating Evolution, Acquisition, and Processing*. Cambridge, MA: MIT Press.

Clark, A. (2015). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.

Conant, R. C., & Ross Ashby, W. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, *1*(2), 89–97. doi: 10.1080/00207727008920220

Constant, A., Ramstead, M. J., Veissiere, S. P., Campbell, J. O., & Friston, K. J. (2018). A variational approach to niche construction. *Journal of the Royal Society Interface*, *15*(141), 20170685. doi: 10.1098/rsif.2017.0685

Constant, A., Ramstead, M. J., Veissière, S. P., & Friston, K. J. (2019). Regimes of expectations: An active inference model of social conformity and human decision making. *Frontiers in Psychology*, *10*, 679. doi: 10.3389/fpsyg.2019.00679

Deacon, T. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W.W. Norton.

Deacon, T. (2011a). *Incomplete Nature: How Mind Emerged from Matter*. New York: W. W. Norton.

Deacon, T. (2011b). The symbol concept. In K. R. Gibson & M. Tallerman (Eds.), *The Oxford Handbook of Language Evolution* (pp. 393–405). Berkeley, CA: University of California Press.

Egan, F. (2019). The nature and function of content in computational models. In M. Sprevak & M. Colombo (Eds.), *The Routledge Handbook of the Computational Mind* (pp. 247–258). London: Routledge.

Eliasmith, C. (2005). A new perspective on representational problems. *Journal of Cognitive Science*, *6*, 97–123.

Fodor, J. A. (1975). *The Language of Thought*, *Vol. 5.* Cambridge, MA: Harvard University Press.

Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, *4*(11), e1000211. doi: 10.1371/journal.pcbi.1000211

Friston, K. & Buzsaki, G. (2016). The functional anatomy of time: What and when in the brain. *Trends in Cognitive Sciences*, *20*(7), 500–511. doi: 10.1016/j.tics.2016.05.001

Friston, K., Parr, T., Yufik, Y., Sajid, N., Price, C. J., & Holmes, E. (2020). Generative models, language and active inference. *PsyArXiv Preprints*. doi: 10.31234/osf.io/4j2k6

Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. doi: 10.1038/nrn2787

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

Friston, K. J. (2020). A free energy principle for a particular physics. ArXiv. 1906.10184v1

Friston, K. J., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: A process theory. *Neural Computation*, *29*(1), 1–49. doi: 10.1162/NECO_a_00912

Friston, K. J., & Frith, C. D. (2015a). Active inference, communication and hermeneutics. *Cortex*, *68*, 129–143. doi: 10.1016/j.cortex.2015.03.025

Friston, K. J., & Frith, C. (2015b). A duet for one. *Consciousness and Cognition*, *36*, 390–405. doi: 10.1016/j.concog.2014.12.003

Friston, K. J., Rosch, R., Parr, T., Price, C., & Bowman, H. (2018). Deep temporal models and active inference. *Neuroscience & Biobehavioral Reviews*, *90*, 486–501. doi: 10.1016/j.neubiorev.2017.04.009

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. doi: 10.3390/e22050516

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Hove: Psychology Press.

Hipólito, I., Ramstead, M. J., Convertino, L., Bhat, A., Friston, K., & Parr, T. (2021). Markov blankets in the brain. *Neuroscience & Biobehavioral Reviews*, *125*, 88–97. doi: 10.1016/j.neubiorev.2021.02.003

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

Ingold, T. (2000). *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill.* New York: Routledge.

Jorna, R. J. J. M. (2006). Neurosemiotics. In M.D. Binder, N. Hirokawa, & U. Windhorst (Eds.), *Encyclopedic Reference of Neuroscience*. Berlin: Springer.

Khani, F., Goodman, N. D., & Liang, P. (2018). Planning, inference and pragmatics in sequential language games. *Transactions of the Association for Computational Linguistics*, *6*, 543–555. doi: 10.1162/tacl_a_00037

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). doi: 10.1098/rsif.2017.0792

Kleinschmidt, D. F. & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148–203. doi: 10.1037/a0038695

Kull, K., Deacon, T., Emmeche, C., Hoffmeyer, J., & Stjernfelt, F. (2011). Theses on biosemiotics: Prolegomena to a theoretical biology. In C. Emmeche & K. Kull (Eds.), *Towards a Semiotic Biology: Life is the Action of Signs* (pp. 25–41). Singapore: Imperial College Press.

Luuk, E. & Luuk, H. (2012). Natural language – no infinity and probably no recursion. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *34*.

MacKay, D. J. C. & Peto, L. C. B. (1995). A hierarchical Dirichlet language model. *Natural Language Engineering*, *1*(3), 289–308. doi: 10.1017/S1351324900000218

Mead, G. H. (1934). *Mind, Self and Society*. Chicago: University of Chicago Press.

Parr, T., Da Costa, L., & Friston, K. (2020). Markov blankets, information geometry and stochastic thermodynamics. *Philosophical Transactions of the Royal Society A*, *378*(2164), 20190159. doi: 10.1098/rsta.2019.0159

Parr, T., & Friston, K. J. (2018). The anatomy of inference: Generative models and brain structure. *Frontiers in Computational Neuroscience*, *12*(90). doi: 10.3389/fncom.2018.00090

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

Peirce, C. S. (1902). Logic as semiotic: The theory of signs. *Philosophical writings of Peirce*. Mineola, NY: Dover.

Pietarinen, A. V. J. (2012). Peirce and Deacon on the meaning and evolution of language. In T. Schilhab, F. Stjernfelt, & T. Deacon, *The Symbolic Species Evolved* (pp. 65–80). Dordrecht: Springer.

Pietarinen, A. V. & Beni, M. D. (2021). Active inference and abduction. *Biosemiotics*, *14*, 499–517. doi: 10.1007/s12304-021-09432-0

Ramsay, D. S. & Woods, S. C. (2014). Clarifying the roles of homeostasis and allostasis in physiological regulation. *Psychological Review*, *121*(2), 225–247. doi: 10.1037/a0035942

Ramstead, M. J. (2015). Naturalizing what? Varieties of naturalism and transcendental phenomenology. *Phenomenology and the Cognitive Sciences*, *14*(4), 929–971. doi: 10.1007/s11097-014-9385-8

Ramstead, M. J., Badcock, P. B., & Friston, K. J. (2018). Answering Schrödinger's question: A free-energy formulation. *Physics of Life Reviews*, *24*, 1–16. doi: 10.1007/s11097-014-9385-8

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

Ramstead, M. J., Kirchhoff, M. D., Constant, A., & Friston, K. J. (2019c). Multiscale integration: Beyond internalism and externalism. *Synthese*, *198*, 41–70. doi: 10.1007/s11229-019-02115-x

Ramstead, M. J., Kirchhoff, M. D., & Friston, K. J. (2019a). A tale of two densities: Active inference is enactive inference. *Adaptive Behavior*, *28*(4), 225–239. doi: 10.1177/1059712319862774

Ramstead, M. J., Friston, K. J. & Hipólito, I. (2020). Is the free-energy principle a formal theory of semantics? From variational density dynamics to neural and phenotypic representations. *Entropy*, 22(8), 889, doi: 10.3390/e22080889

Ramstead, M. J., Veissière, S. P., & Kirmayer, L. J. (2016). Cultural affordances: Scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in Psychology*, 7. doi: 10.3389/fpsyg.2016.01090

Roy, D. (2005). Semiotic schemas: A framework for grounding language in action and perception. *Artificial Intelligence*, *167*(1–2), 170–205. doi: 10.1016/j.artint.2005.04.007

Sterelny, K. (2010). Minds: extended or scaffolded? *Phenomenology and the Cognitive Sciences, 9*(4), 465–481. doi: 10.1007/s11097-010-9174-y

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

Seifert, U. (2012). Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, *75*(12). doi: 10.1088/0034-4885/75/12/126001

Seth, A. K. (2015). Inference to the best prediction. In T. K. Metzinger & J. M. Windt (Eds.), *Open MIND*. Frankfurt am Main: MIND Group.

Seth, A. K. & Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708). doi: 10.1098/rstb.2016.0007

Short, T. L. (2007). *Peirce's Theory of Signs*. Cambridge: Cambridge University Press.

Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A. E., Paliwal, S., Gard, T., Tittgemeyer, M., Fleming, S. M., Haker, H., Seth, A. K., Petzschner, F. H. (2016). Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression. *Frontiers in Human Neuroscience*, *10*, 550. doi: 10.3389/fnhum.2016.00550

Sterling, P. & Eyer, J. (1988). Allostasis: A new paradigm to explain arousal pathology. In S. Fisher & J. Reason (Eds.), *Handbook of Life Stress, Cognition and Health* (pp. 629–649). New York: John Wiley & Sons.

Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006). Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, *101*(476), 1566–1581. doi: 10.1198/016214506000000302

Thompson, E. (2010). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.

Tschantz, A., Seth, A. K., & Buckley, C. L. (2019). Learning action-oriented models through active inference. bioRxiv. doi: 10.1101/764969

Varela, F. J. (1996). *Neurophenomenology: A Methodological Remedy for the Hard Problem*. Exeter: Imprint Academic.

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

Vasil, J., Badcock, P. B., Constant, A., Friston, K., & Ramstead, M. J. (2019). A world unto itself: Human communication as active inference. arXiv, arXiv: 1906.10538

Veissière, S. P., Constant, A., Ramstead, M. J., Friston, K. J., & Kirmayer, L. J. (2020). Thinking through other minds: A variational approach to cognition and culture. *Behavioral and Brain Sciences*, *43*, e90. doi: 10.1017/S0140525X19001213

von Uexküll, J. (1982). The theory of meaning. *Semiotica*, *42*(1), 25–79. doi: 10.1515/semi.1982.42.1.25

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*(2), 158–177. doi: 10.1037/h0074428