3 The High Road to Active Inference

Survival machines that can simulate the future are one jump ahead of survival machines who can only learn on the basis of overt trial and error. The trouble with overt trial is that it takes time and energy. The trouble with overt error is that it is often fatal. Simulation is both safer and faster.

—Richard Dawkins

3.1 Introduction

In chapter 2, we motivated the introduction of free energy as a means of performing approximate Bayesian inference (i.e., the low road to Active Inference). Here, we introduce free energy from another perspective, that of the high road, which inverts that reasoning: it starts from first principles in statistical physics and the central imperative that organisms must maintain their existence—that is, avoid surprising states—and then introduces the minimization of free energy as a computationally tractable solution to this problem. The chapter discloses the formal equivalence between the minimization of variational free energy and the maximization of model evidence (or self-evidencing) in approximate Bayesian inference, revealing a connection between free energy and Bayesian perspectives on adaptive systems. Finally, it discusses how Active Inference provides a novel first principle perspective to understand (optimal) behavior.

Active Inference is a theory of how living organisms maintain their existence by minimizing surprise—or a tractable proxy to surprise, *variational free energy*—via perception and action. By starting from first principles, it advances a novel belief-based scheme to understand behavior and cognition, which has numerous empirical implications.

The high road to Active Inference starts from the premise that, to survive, any living organism has to maintain itself in a suitable set of preferred states, while avoiding other, dis-preferred states of the environment. These preferred states are first and foremost defined by niche-specific evolutionary adaptations. However, as we will see later, in advanced organisms these can also extend to learned cognitive goals. For example, to survive, a fish has to stay in a comfort zone that corresponds to a small subset of all the possible states of the universe: it has to stay in water. Similarly, a human has to ensure that their internal states (e.g., physiological variables like body temperature and heart rate) always remain within acceptable ranges—otherwise they will die (or more precisely will become something else, such as a corpse). This acceptable range or comfort zone stipulatively defines the characteristic states something has to be in to be that thing.

Living organisms resolve this fundamental biological problem by exerting *active control* over their states (e.g., of body temperature) at many levels, which range from automatic regulatory mechanisms such as sweating (physiology) to cognitive mechanisms such as buying and consuming a drink (psychology) to cultural practices such as distributing air conditioning systems (social sciences).

From a more formal perspective, Active Inference casts the biological problem of—or explanation for—survival as surprise minimization. This formulation rests on a technical definition of *surprising states* from information theory—essentially, surprising states index those outside the comfort zone of living organisms. It then proposes free energy minimization as a practical and biologically grounded way for organisms or adaptive systems to minimize the surprise of sensory encounters.

3.2   Markov Blankets

An important precondition for any adaptive system is that it must enjoy some separation and autonomy from the environment—without which it would simply dissipate, dissolve, and thereby succumb to environmental dynamics. In the absence of this separation, there would be no surprise to minimize; there must be something to be surprised and something to be surprised about. In other words, there are at least two things—system and environment—and these can be disambiguated from one another. A formal way to express a separation between a system and the rest of the environment is the statistical construct of a *Markov blanket* (Pearl 1988); see box 3.1.

Box 3.1

Markov blankets

A Markov blanket is an important recurring concept in this book (Friston 2019a, Kirchhoff et al. 2018, Palacios et al. 2020). Technically, a blanket (b) is defined as follows:

$$\mu \perp x \,|\, b \Leftrightarrow p(\mu, x | b) = p(\mu | b) p(x | b)$$

This says (in two different but equivalent ways) that a variable μ is conditionally independent of a variable x if b is known. In other words, if we know b, knowing x would give us no additional information about μ. A common example of this is a Markov chain, where the past causes the present causes the future. In this scenario, the past may only influence the future via the present. This means no additional information about the future is gained by finding out about the past (assuming we know the present).

To identify a Markov blanket in a system wherein we know the conditional dependencies, we can follow a simple rule. The blanket for a given variable comprises its *parents* (the variables it depends on), its *children* (the variables that depend on it) and, in some settings, the other *parents of its children*.

In brief, a Markov blanket is the set of variables that mediate all (statistical) interactions between a system and its environment. Figure 3.1 illustrates an interpretation of a Markov blanket in a dynamic setting. Here the conditional independences have been supplemented with dynamical constraints, so that the flows do not depend upon states on the opposite side of the blanket.

The Markov blanket in figure 3.1 distinguishes states *internal* to the adaptive system (i.e., brain activity) from *external* states of the environment. Furthermore, it identifies two additional states, labeled *sensory states* and *active states*, which form the blanket that (statistically) separates internal and external states. Statistical separation means that if we knew about the active and sensory states, the external states would offer no additional information about internal states (and vice versa). In a dynamical setting, this is often interpreted as saying internal states cannot directly change external states but can do so vicariously by changing *active states*. Similarly, external states cannot directly change internal states but can do so indirectly by changing *sensory states*.

This is a restatement of the classical action-perception cycle, wherein an adaptive system and its environment can interact (only) through actions and observations, respectively. This reformulation has two main benefits.
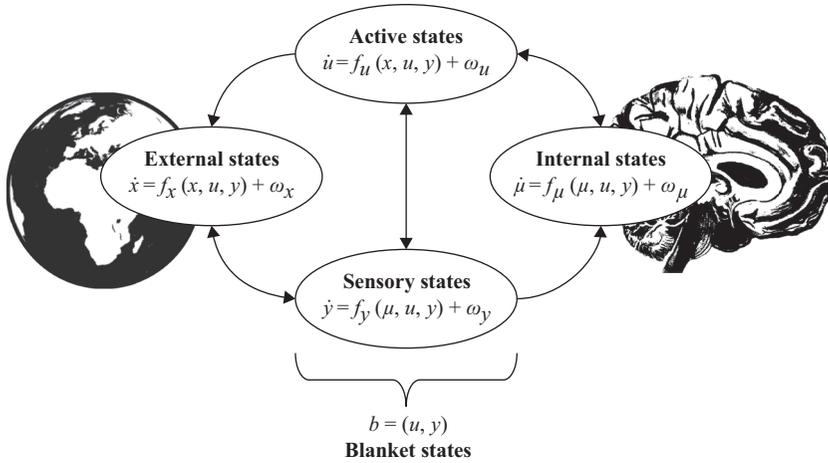
Figure 3.1

A dynamic Markov blanket, which separates an adaptive system (here, the brain) from the environment. The dynamics of each set of states are determined by a deterministic flow specified as a function (f) giving the average rate of change and additional stochastic (random) fluctuations (ω). The arrows indicate the direction of influence of each variable over the rates of change of other variables (technically, the nonzero elements of the associated Jacobians). This is just one example; one can use a Markov blanket to separate an entire organism from the environment or nest multiple Markov blankets within one another. For example, brains, organisms, dyads, and communities can be conceived in terms of different Markov blankets that are nested within one another (see Friston 2019a; Parr, Da Costa, and Friston 2020 for a formal treatment). Confusingly, different fields use different notations for the variables; sometimes, sensory states are denoted s, external states η, and active states a. Here we have chosen variables for consistency with the other chapters in this book.

First, it formalizes the fact that an adaptive system's internal states are autonomous from environmental dynamics and can therefore resist their influences. Second, it scaffolds the way in which adaptive systems minimize their surprise: it highlights the internal, sensory, and active states they have access to. Specifically, surprise is defined in relation to sensory states, while internal and active state dynamics are the means by which the surprise of sensory states may be minimized.

The key point to notice here is that the internal states of an adaptive system bear a formal relation to external states. This is due to a kind of symmetry across the Markov blanket as both influence and are influenced by blanket states. A consequence of this is that we can construct conditional

probability distributions for the internal and external states, given the blanket states. Because these are conditioned on the same blanket states, we can associate pairs of expected internal and external states with one another. In other words, on average, the internal and external states acquire a kind of (generalized) synchrony—just as we might anticipate on attaching a pendulum to each end of a wooden beam. Over time, as they synchronize, each pendulum becomes predictive of the other through the vicarious influence of the beam (Huygens 1673). Figure 3.2 offers a graphical intuition for this relationship. This means that if we can write down independent
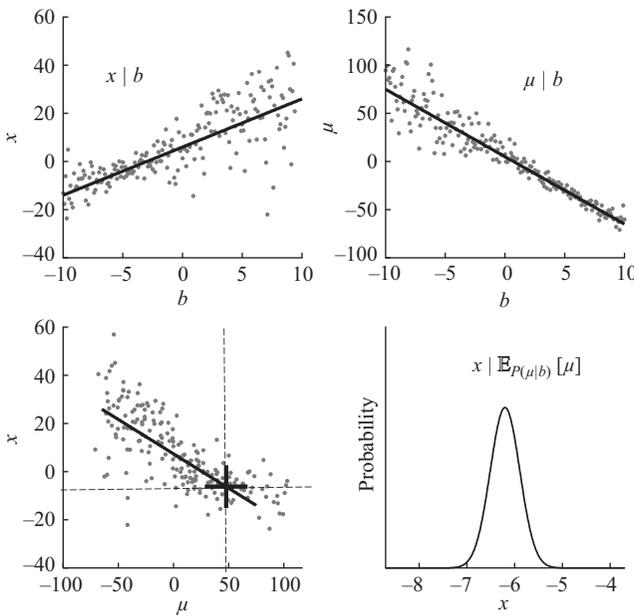

Figure 3.2

Association between average internal states of a Markov blanket and distributions of external states. *Top:* Assuming a linear Gaussian form for the conditional probabilities, these plots show samples from the conditional distribution over external and internal states, respectively, given blanket states. The thick black lines indicate the average of these variables given the associated blanket state. *Bottom left:* The same data are plotted to illustrate the synchronization of internal and external states afforded by sharing a Markov blanket—here, an inverse synchronization. The dashed lines and black cross illustrate that if we knew the average internal state (vertical line), we could identify the average external state (horizontal line) and the spread around this point. *Bottom right:* We can associate the average internal state with a distribution over the external state.

distributions over external and internal states given their Markov blanket, the two states become informative about one another via this blanket.

This synchrony gives internal states the appearance of representing (or *modeling*) external states—which links back to the idea of surprise minimization introduced in chapter 2. This is because surprise depends on an internal model of how sensory data are generated. To recap, minimizing the surprise (negative log probability) of sensory observations becomes identical to maximizing the evidence (marginal likelihood) for the model, which is just the probability of sensory observations under that model. This notion of surprise minimization can be understood from two equivalent—Bayesian and free energy—perspectives, which we discuss next.

3.3 Surprise Minimization and Self-Evidencing

Under a Bayesian perspective, an agent with a Markov blanket appears to *model* the external environment in the sense that internal states correspond (on average) to a probabilistic representation—an approximate *posterior belief*—of external states of the system (figure 3.2). The dynamics of internal states correspond to a form of (approximate) Bayesian inference of external states, as their motion changes the associated probability distribution, which is afforded by an implicit generative model of how sensations (or *sensory states* in the Markov blanket jargon) are generated. If we reinstate the notion of an agent as constituted by internal and blanket states, we can talk about an agent's generative model.

Importantly, the agent's generative model cannot simply mimic external dynamics (otherwise the agent would simply follow external dissipative dynamics). Rather, the model must also specify the preferred conditions for the agent's existence, or the regions of states that the agent has to visit to maintain its existence, or satisfy the criteria for its existence in terms of occupying characteristic states. These preferred states (or observations) can be specified as the *priors* of the model—which implies that the model implicitly assumes that its preferred (prior) sensations are more likely to occur (i.e., are less surprising) if it satisfies the criteria for existence. This means it has an implicit *optimism bias*. This optimism bias is necessary for the agent to go beyond the mere duplication of external dynamics to prescribe *active states* that underwrite its preferred or characteristic states.

Under this formulation, one can cast optimal behavior (with respect to prior preferences) as the *maximization of model evidence by perception and*

action. Indeed, model evidence summarizes how well the generative model fits or explains sensations. A good fit indicates that the model successfully accounts for its sensations (this is the *descriptive* side of inference); at the same time, it realizes its preferred sensations, given that they are less surprising (this is the *prescriptive* side of the inference). Such good fit is a guarantee of surprise minimization, as maximizing model evidence $P(y)$ is mathematically equivalent to minimizing surprise: $\Im(y) = -\ln P(y)$.

A way to reformulate the above arguments more succinctly consists in saying that any adaptive system engages in "self-evidencing" (Hohwy 2016). *Self-evidencing* here means acting to garner sensory data consistent with (i.e., that affords evidence to) an internal model, hence maximizing model evidence.

3.3.1   Surprise Minimization as a Hamiltonian Principle of Least Action

In the preceding sections, we have asserted that surprise must be minimized but have not detailed why this is. Although the details of the underlying physics of self-evidencing are outside the scope of this book (see Friston 2019b for details), we here provide a brief overview of the principles. These are underwritten by the idea that biological creatures—with Markov blankets—persist over time, resisting the dispersive effects of environmental fluctuations. The persistence of a Markov blanket implies that the distribution of blanket states remains constant over time. Simply put, this means that any deviation of sensory (or active) states from regions that are highly probable under this distribution must be corrected by the average flow of states (which is just the deterministic part of the flow in figure 3.1). Expressing this as a physicist might, stochastic (random) systems at steady state engage in dynamics that (on average) descend an energy function (or Hamiltonian) that is interpretable as a negative log evidence or surprise. This is like a ball rolling down a hill from high gravitational potential energy at the top of the hill to low energy in a basin. See figure 3.3.

For the system shown on the left of figure 3.3, every time a fluctuation causes a move to a less probable state, this is corrected by a move up the probability gradient, such that the system occupies probability-dense regions a greater proportion of the time. The key insight here is that this system maintains sensory states within a narrow range by minimizing surprise (on average)—in contrast to the system on the right, for which surprise grows indefinitely.

Surprise minimization permits living organisms to (temporarily) resist the second law of thermodynamics, which states that entropy—or the
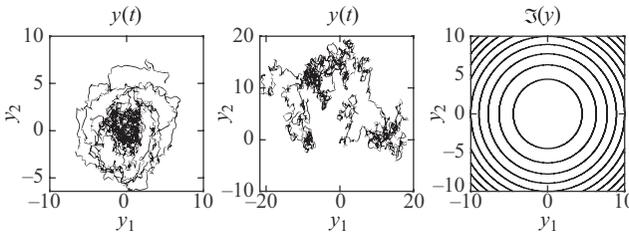
Figure 3.3

Left: Path taken by a 2-dimensional random dynamical system with a (nonequilibrium[1]) steady state. This can be interpreted as minimizing its surprise, which is shown in the contour plot on the right. *Right:* The center is the least surprising region; the circles moving away from the center represent progressively more surprising regions. *Middle:* In contrast, this plot shows the trajectory of a system starting in the same place $(5, 5)$, with random fluctuations of the same amplitude, whose dynamics bear no relation to surprise. Not only does it enter more surprising regions of space; it also fails to achieve any sort of steady state, dissipating in an unconstrained fashion over time. The scope of Active Inference is restricted to systems like that on the left—which counter random fluctuations with their average flow and thereby retain their form over time.

dispersion of systemic states—always grows. This is because, on average, entropy is the long-term average of surprise and, on average, the maximization of a log probability of observations is equivalent to minimization of (Shannon) entropy:[2]

$$H[P(y)] = \mathbb{E}_{P(y)}[\mathfrak{I}(y)] = -\mathbb{E}_{P(y)}[\ln P(y)] \tag{3.1}$$

Ensuring that a small proportion of sensory states is occupied with high probability is equivalent to maintaining a particular entropy. This is a defining characteristic of self-organizing systems, as long recognized by cybernetic theories.

From a physiologist's perspective, surprise minimization formalizes the idea of homeostasis. As a sensor value leaves its optimal range, negative feedback mechanisms kick in that reverse these deviations. From a control perspective, we can interpret optimal behavior in relation to some desired steady state probability density. In other words, if we define a distribution of preferred outcomes, optimal behavior will involve evolution of the system toward—and maintenance of—that distribution.

As we saw in chapter 2, free energy is an upper bound on surprise, suggesting that optimal behavior can be obtained by minimizing free energy

in the face of random fluctuations. Recall that the difference between free energy and surprise is the divergence between an exact posterior probability (i.e., the distribution of external states given blanket states) and an approximate posterior probability (i.e., the distribution over external states given average internal states). As such, the motion of internal states can be thought of as minimizing the divergence, which then enables active states, on average, to minimize the surprise accompanying sensory states. In other words, the optimal behavior resulting from free energy minimization is the one that is least surprising and follows a path of least Action[3] from the current state to the desired state—that is, the Hamiltonian principle of least Action applied to behavior.

Figure 3.3 shows a very simple example of a system equipped with a random attractor. This is analogous to a thermostat, which (in cybernetic parlance) has a single set-point and cannot learn or plan. Active Inference aims to use the same explanatory apparatus to cover much more complex and adaptive systems. Here, the difference between simplest and more complex systems can be reduced to the different shapes of their attractors—from fixed points to increasingly more complex and itinerant dynamics. From this perspective, one can understand living organisms as constantly seeking a compromise between excessive stability and excessive dispersion—and Active Inference aims to explain how such compromise is achieved.

3.4 Relations between Inference, Cognition, and Stochastic Dynamics

The physicist E. T. Jaynes famously argued that inference, information theory, and statistical physics are different perspectives on the same thing (Jaynes 1957). In the previous sections, we discussed how Bayesian and statistical physics perspectives offer two equivalent ways to understand surprise minimization and optimal behavior—effectively adding a form of cognition to Jaynes's triad. This equivalence between various schools of thought is appealing but can be confusing to those who are not familiar with the respective formalisms, where many different words are used to refer to the same quantities. To help demystify this, in this section we elaborate on the main equivalences between Bayesian and statistical physics perspectives and their cognitive interpretations; see table 3.1 for a summary and box 3.2.

Table 3.1
Statistical physics, Bayesian inference, and information theory—and their cognitive interpretations

Statistical physics	Bayesian inference and information theory	Cognitive interpretation
Minimize variational free energy	Maximize model evidence (or marginal likelihood); minimize surprisal (or self-information)	Perception and action
Minimize expected free energy; Hamiltonian principle of least Action	Infer the most likely (or less surprising) course of action	Planning as inference
Attain nonequilibrium steady-state	Perform approximate Bayesian inference	Self-evidencing
Gradient flows on energy functions; gradient descent on free energy	Gradient ascent on model evidence; gradient descent on surprisal	Neuronal dynamics

Box 3.2
Free energy in statistical physics and Active Inference

The notion of free energy is widely used in statistical physics to characterize (for example) thermodynamic systems. Although Active Inference uses exactly the same equations, it applies them to characterize the *belief state* of an agent (in relation to a generative model). Hence, when we talk of an Active Inference agent minimizing its (variational) free energy, we are referring to processes that change its belief state, not (for example) the particles of its body. To avoid misunderstandings, we use the term *variational free energy*, hence adopting a terminology that is more common in machine learning. Another more subtle point is that the concept of free energy is often used in the context of equilibrium statistical thermodynamics. Active Inference targets living organisms—or nonequilibrium steady state systems that are open—that feature continuous, reciprocal exchanges with the environment. This is an exciting novel field (Friston 2019a).

3.4.1 Variational Free Energy, Model Evidence, and Surprise

A first important equivalence is between the *maximization of model evidence* (or marginal likelihood) in Bayesian inference and the *minimization of variational free energy*—both of which minimize surprise. This equivalence becomes evident when one appeals to a specific approximate solution to intractable problems of inference—variational inference. Variational inference recasts the inference problem as an optimization problem by minimizing free energy. The minimum of the free energy is the point at which the approximation of the exact solution is at its best. Expressing this formally sheds light on the relations between the three quantities:

$$\underbrace{\Im(y|m)}_{\text{Surprise}} = -\ln \underbrace{P(y|m)}_{\text{Model evidence}} \leq \underbrace{D_{KL}\big[Q(x)\,\|\,P(x|y,m)\big] - \ln P(y|m)}_{\text{Variational free energy}} \tag{3.2}$$

In equation 3.2, unlike in chapter 2, we have explicitly conditioned all quantities on a model, *m*, to emphasize that these depend on the model we have (or are) about how *y* are generated, and the quantities will vary if different models are used. The equivalence of these quantities raises the question as to why it is useful to distinguish between them. The main reason is that, unlike model evidence, variational free energy can be minimized efficiently.

Recall from chapter 2 that the variational free energy is only exactly equivalent to the negative model evidence or surprise when the KL-Divergence term becomes zero. This is not always possible, but this can be made close to zero. Hence, in the process of finding better and better values for $Q(x)$, variational free energy also approximates surprise more closely. We have said this a few times already because it is important to emphasize the central relationship between free energy and surprise that is the foundation of this book. Specifically, free energy is an *upper bound* on surprise. It can be the same as or greater than surprise—where what is *greater than* is quantified by the KL-Divergence.

An interesting aspect of this is that any system minimizing its surprise, including the very simple system in figure 3.2, is also minimizing a free energy, where the $Q(x)$ is always set to be equal to the exact posterior probability—that is, setting the KL-Divergence to be zero. One perspective on the difference between cognitive and noncognitive systems is that the latter always have a zero KL-Divergence, while cognitive systems must go through the (perceptual) process of minimizing this term before their actions are guaranteed to minimize surprise. Note that minimizing the divergence is the only thing that perception can do. This places a great

deal of emphasis on the motion of internal states, such that the distribution they parameterize (figure 3.2) is as close to the exact posterior as possible. However, perception cannot minimize the second (evidence) component of variational free energy that corresponds to the actual surprise, because it cannot change the sensations that have been gathered. Only by acting in ways that change sensations can an agent minimize the second (evidence) component of variational free energy and resolve its surprise—or, equivalently, maximize its model evidence. This places emphasis on the motion of active states, given internal states, in self-evidencing.

An example helps in illustrating this point. Imagine that your generative model predicts a distribution of glucose levels in your blood given levels of hunger, with relatively high versus low glucose levels relating to satiation and hunger, respectively. In addition, imagine this model ascribes a higher prior probability to satiation and therefore to relatively high glucose levels—making low glucose levels surprising. Imagine you are initially uncertain about your hunger levels and sense low blood glucose. Perception leads to the inference that you are hungry and the experience of hunger—closing the KL-Divergence. However, perception cannot go further than that to reduce your surprise—and the discrepancy between the high level of glucose that you expect a priori and the low level of glucose that you sense—because it cannot act on your sensations (low glucose) or their causes (physiology). You can only minimize your surprise by acting to change (the hidden source of) the sensations you gather—for example, by eating a dessert.

In sum, perception can minimize variational free energy by reducing the discrepancy between approximate and true posterior but cannot go further in minimizing surprise. The next step of surprise minimization entails changing the sensations one gathers by acting, which is where inference goes beyond perception and becomes *active*.

3.4.2 Expected Free Energy and Inference of the Most Likely Trajectory

Another important equivalence is between the minimization of *expected free energy* and inferring the *most likely course of action*, or *policy*. This goes beyond specifying the least surprising part of state-space and deals with how surprising alternative routes to that part or location may be. These alternative paths are expressed in terms of policies, which are essentially trajectories across states. Importantly, in Active Inference the log probability of a policy is set proportional to the expected free energy if that policy

was pursued. This implies that the most probable or least surprising path is (set to be) the one that minimizes expected free energy. This formulation is equivalent to the way *Action* is defined in physics, where it scores the probability of a path by an integral (or sum) of an energy. While a physical system may pursue a space of hypothetical trajectories, the path it actually follows is the one for which Action is minimized—that is, Hamilton's principle of least Action. This analogy between Active Inference and Hamilton's principle of least Action is unpacked in the next section.

3.5 Active Inference: A Novel Foundation to Understand Behavior and Cognition

In fields like optimal control, reinforcement learning, and economics, the optimization of behavior results from a *value function* of states, following Bellman's equation (Sutton and Barto 1998). Essentially, each state (or state-action pair) is assigned a value, which represents how good a state is for an agent to be in. The value of states (or state-action pairs) is usually learned by trial and error, by counting how many times—and after how much time—one obtains reward by starting from those states. Behavior consists in optimizing reward acquisition by reaching high-valued states, hence capitalizing on learning history.

In contrast, in Active Inference, behavior is the result of inference and its optimization is a function of beliefs. This formulation unites notions of (prior) belief and preference. As discussed above, using the notion of expected free energy amounts to endowing the agent with an implicit prior belief that it will realize its preferences. Hence, the agent's preference for a course of action becomes simply a belief about what it expects to do, and to encounter, in the future—or a belief about future trajectories of states that it will visit. This replaces the notion of value with the notion of (prior) belief. This is an apparently strange move, if one has a background in reinforcement learning (where value and belief are separated) or Bayesian statistics (where belief does not entail any value). However, it is a powerful move, for at least three reasons.

First, it automatically entails a self-consistent process model of purposive (or teleological) behavior, which is akin to cybernetic formulations. If we endow an Active Inference agent with some prior preference, then it will act to realize such preferences—because this is the only course of action

consistent with its prior belief that it will act to fulfill its expectations. Note that the resulting (preferred) course of action, or policy, is directly measurable in experimental settings, whereas a value function or prior belief needs to be inferred and hence is a more indirect, if not tautological, measure.

Second, casting behavior as a functional of beliefs (probability distributions) automatically entails notions such as *degree of belief* and *uncertainty*. These notions undergird important facets of adaptive action but are not directly available in the Bellman formulation. By the same token, this formulation gives more flexibility in modeling sequential dynamics and itinerant behaviors, which are harder to model in terms of a value function of states (Friston, Daunizeau, and Kiebel 2009).

Third, in this formulation, optimal behavior comes to follow a Hamiltonian principle of least Action in statistical physics. Indeed, Active Inference goes one step further toward the idea that behavior is a function of beliefs: it also assumes that it becomes an energy function—and the most likely course of action of an Active Inference agent is the one that minimizes free energy. A profound consequence is that living organisms behave according to Hamilton's principle of least Action: they follow a path of least resistance until they reach a steady state (or a trajectory of states), as exemplified by the behavior of a random dynamical system (shown in figure 3.3). This is a fundamental assumption that distinguishes Active Inference from alternative theories of behavior and cognition based on the Bellman formulation.

It is worth briefly outlining what we mean by drawing analogies between Hamiltonian physics and Active Inference. This is intended on three levels. The first is that the advance offered by Active Inference to the behavioral and life sciences is comparable to the advance Lagrangian[4] and Hamiltonian formulations offered to Newton's accounts of mechanics. While Newtonian mechanics were originally formulated in terms of differential equations—including Newton's famous third law expressing the proportionality between acceleration and force—a complementary perspective on mechanics was offered by considering what is conserved by dynamical systems. Newtonian dynamics can then be derived from these conservation laws. These offer a perspective on which to base further theoretical advances, and they form the basis for parts of stochastic, relativistic, and quantum physics. Analogously, Active Inference reformulates the sorts of neuronal and behavioral dynamics that might previously have been built up from a series of differential equations by specifying the quantity—free energy—from

which these dynamics may be derived. Just as different sorts of Hamiltonians lead to different types of physics, free energies based on different generative models lead to different neuronal and behavioral dynamics.

The second point of connection between Hamiltonian physics and Active Inference arises from a more direct association between a Hamiltonian and probability measures. The idea here is to associate the conserved Hamiltonian with the energy of the system. Remember that the quantities we have referred to as *energies* so far (here and in chapter 2) have all had the form of a negative log probability. This reflects an interpretation of energy as simply a measure of the improbability of any given configuration of a system. On this view, conservation of energy and of probability are equivalent laws. As dissipative systems—coupled to external states via a Markov blanket—move to states of low energy or high probability, we can directly associate the energy or Hamiltonian with surprise. As such, Active Inference *is* Hamiltonian physics applied to a certain kind of system (systems that feature a Markov blanket).

The third association between these formulations is the variational calculus that underwrites the association between energies and dynamics. This is most apparent when Hamiltonian physics is expressed as a principle of least Action, where *Action* refers to the integral of a Lagrangian over a path. Crucially, this Action is a functional of a path. Here, a path is a function of time whose output is the position and velocity of a particle on that path at that time. The path followed by a (deterministic) particle minimizes this Action. Similarly, Active Inference is predicated on the idea that beliefs (themselves functions of hidden states) must minimize a free energy functional. The key point of contact here is that in both cases, functions (paths or beliefs) must be optimized in relation to functionals (Action or free energy, respectively). This places both in the context of variational calculus, which is a branch of mathematics dedicated to finding extrema of functionals. In physics, this leads to the Euler-Lagrange equations. In Active Inference, we arrive at variational inference procedures.

3.6 Models, Policies, and Trajectories

In section 3.2, we highlighted that the scope of Active Inference pertains to those systems that enjoy some separation from their environment and saw that this translates into the presence of a Markov blanket. In section 3.3,

we highlighted that the persistence of this blanket requires dynamics that (on average) minimize the surprise of (sensory) states. As this may be interpreted as self-evidencing, we arrive at the conclusion that behavior is determined by a steady-state distribution that can be interpreted as a generative model of how (sensory) data are generated.

This tells us something very important. Each generative model should be associated with different sorts of behavior. As such, different sorts of behavior may be accounted for by specifying different generative models—and implicitly what that system would find surprising. Furthermore, different kinds of generative model may correspond to adaptive or cognitive creatures having various levels of complexity (Corcoran et al. 2020). Very simple generative models of the sort driving the dynamics in figure 3.3 offer a minimal sort of cognition, as they cannot entertain the possibility of alternative (or counterfactual) trajectories. Further, these models are shallow, in the sense that they afford inference at just one timescale. In contrast, *hierarchical* generative models afford inference at multiple timescales. In hierarchical or deep models, the dynamics at higher hierarchical levels generally encode things that change more slowly (e.g., the sentence I am reading) and that contextualize things that change faster (e.g., the word I am reading), which are represented at lower hierarchical levels (Kiebel et al. 2008; Friston, Parr, and de Vries 2017).

What do we need to include in a model to derive more complex behaviors of the sort we would associate with agency and sentient systems? One answer to this is the capacity to model alternative futures, or different ways in which events might play out—and to select among them. In turn, considering possible futures requires a generative model that has some temporal depth and explicitly represents the consequences of actions. Working this into the model will ensure behavior that conforms to the most likely of these futures. The (counterfactual) capacity to entertain these alternatives may be what separates the steady state associated with sentient systems from simpler creatures. When alternative futures pertain to things over which we have control, we refer to these as policies or plans. As we saw in chapter 2, one way of disambiguating between these plans is to incorporate a prior belief into a model that says that those policies with the lowest expected free energy are the most plausible. This offers a way of characterizing a certain kind of system with a Markov blanket at steady state—which seems to correspond well to systems like us.

3.7   Reconciliation of Enactive, Cybernetic, and Predictive Theories under Active Inference

By emphasizing free energy minimization, Active Inference unites and extends three apparently disconnected theoretical perspectives.

First, Active Inference is in keeping with enactive theories of life and cognition, which emphasize the self-organization of behavior and autopoietic interactions with the environment, which ensure that living organisms remain within acceptable bounds (Maturana and Varela 1980). Active Inference provides a formal framework explaining how living organisms manage to resist the dispersion of their states by self-organizing a statistical structure—the Markov blanket—that affords reciprocal exchanges between organism and environment while also separating (and in a sense protecting the integrity of) the organisms' states from external, environmental dynamics.

Second, Active Inference is in keeping with cybernetic theories, which describe behavior as purposive and teleological. *Teleology* means that behavior is internally regulated by a mechanism that continuously tests whether a goal is achieved and, if not, steers corrective actions (Rosenblueth et al. 1943, Wiener 1948, Ashby 1952, G. Miller et al. 1960, Powers 1973). Similarly, Active Inference agents use both perception and action to minimize the discrepancy between preferred and sensed states. Active Inference provides a normative and viable description of the minimization process by specifying that what is actually minimized is a statistical quantity that the agent can measure—variational free energy—which under certain conditions corresponds to a *prediction error*, or the difference between what is expected and what is sensed. This implies a formulation of cybernetic control as a prospective process—which leads us to the next point.

Third, Active Inference is in keeping with theories that describe control as a prospective process that rests on a model of the environment—possibly physically implemented in the brain (Craik 1943). Active Inference assumes that agents use a (generative) model to construct predictions that guide perception and action and to evaluate their future (and counterfactual) action possibilities. This assumption is coherent with the good regulator theorem (Conant and Ashby 1970), which says that any controller should have—or be—a good model of the environment. Active Inference reconciles these model-based perspectives on brain and behavior under a rigorous characterization in terms of (approximate) Bayesian inference and (variational

and expected) free energy minimization. Furthermore, Active Inference is largely coherent with ideomotor theory (Herbart 1825, James 1890, Hoffmann 1993, Hommel et al. 2001), which states that action starts with an imaginative process, and it is a predictive representation (of action consequences) that triggers actions—not a stimulus, like in stimulus-response theory (Skinner 1938). Active Inference casts this idea in an inferential framework, in which an action stems from a belief (about the future); this has a number of implications, such as the fact that in order to trigger action, one has to temporarily attenuate sensory evidence (which would otherwise falsify the belief that triggers action) (H. Brown et al. 2013).

The reconciliation of these frameworks is interesting, as they are often considered at odds. For example, self-organization and teleology are often seen as incompatible in biology. Furthermore, enactive theories tend to de-emphasize representation and control, which is instead a central construct of most theories of model-based inference. Active Inference formalizes autopoietic dynamics of adaptive agents from an unusual angle, which simultaneously considers self-organization and prediction. By connecting different perspectives, Active Inference can potentially help us understand how they illuminate one another.

3.8   Active Inference, from the Emergence of Life to Agency

Active Inference starts from first principles and unfolds them to explain behavior and cognition expressed by the simplest to the most complex forms of adaptive and living systems. In the continuum between simpler and more complex creatures, Active Inference draws a line between those that minimize variational free energy and those that also minimize expected free energy.

Any adaptive system that actively samples sensations to minimize *variational free energy* is (equivalently) an agent that actively gathers evidence for its generative model, aka a *self-evidencing agent* (Hohwy 2016). These systems are able to avoid dissipation, self-regulate, and survive by achieving set-points provided by basic homeostatic processes. These systems can generate complex and diverse forms of behavior and can also have very high fitness levels (as is already apparent in the case of viruses). Some may have hierarchical generative models that permit inferring events that change at different timescales, from faster (at lower hierarchical levels) to slower

(at higher levels)—and hence can develop sophisticated strategies to deal with what they experience. However, these creatures are also fundamentally limited because their generative models lack temporal depth—or the capacity to plan and consider the future explicitly (although they can do so implicitly, for example, as a result of genetic evolution)—and hence they always live in the present.

A generative model endowed with temporal depth opens the door to the minimization of expected free energy—or in psychological terms, planning. In Active Inference, this entails much more than increased adaptivity: it entails at least a primitive form of agency. For an adaptive system, minimizing expected free energy is equivalent to having the (implicit) prior that one is a free energy minimizing agent—but acts to minimize free energy in the future. When this (prior) belief enters the generative model, the adaptive system becomes able to form beliefs about how it should behave in the future and which trajectories it will pursue. In other words, it becomes able to *select among alternative futures* as opposed to simply selecting *how to deal with the sensed present*, as in the simplest agents described above. This temporal depth therefore translates into a psychological depth. To ask about the ways living creatures populate the continuum between the simplest and most complex adaptive systems—and what forms of Active Inference they can express—is an empirical question.

3.9 Summary

The main topics of this chapter can be summarized as follows: Living organisms have to ensure that they only visit their characteristic or preferred states. If one defines these preferred states as expected states, then one can say that living organisms must minimize the surprise of their sensory observations (and maintain an optimal entropy; see box 3.3).

Doing this requires agents to exercise some autonomy from environmental dynamics and to be equipped with a Markov blanket that separates (i.e., expresses a conditional independence between) their internal states and the external states of the environment. Agents within the Markov blanket can engage in reciprocal (action-perception) exchanges with the environment. These exchanges are formally described by the theory of Active Inference, where both perception and action minimize surprise. They can do so by being equipped with a probabilistic generative model of how their

Box 3.3
Entropy minimization and open-ended behavior

Active Inference is based on the premise that living organisms strive to maintain a relative order (or negative entropy), controllability and predictability, despite being immersed in an environment whose natural forces generate continuous fluctuations—and a never-ending threat of entropic erosion. The most basic manifestation of this active pursuance of order is physiological homeostasis, with critical physiological parameters that need to be kept within viable regions. However, minimizing entropy should not be equated with a rigid repertoire of responses (e.g., autonomic homeostatic responses) but rather the opposite, especially in advanced organisms. We can develop open-ended repertoires of novel behaviors to pursue our original homeostatic imperatives—for example, to produce and buy good wine to satisfy thirst and other needs. This is sometimes referred to as "allostasis" (Sterling 2012).

More broadly, we actively pursue some order and controllability per se, without necessary reference to a specific homeostatic imperative—perhaps because preserving order facilitates many such imperatives. We actively carve our ecological niches to render them more predictable and less surprising. This is evident in the ways we construct our physical spaces (e.g., refuges and cities that give shelter from uncontrolled natural forces) and cultural spaces (e.g., societies with laws and deontic norms that give shelter from anarchic social forces). In all these examples, we usually need to accept some short-term increase of entropy or surprise (e.g., when we build something new or shift social stances) to ensure their long-term decrease. This helps us understand how the basic requirement for surprise minimization is not at odds with but rather promotes the epistemic imperatives and novelty-seeking, curious, and exploratory behavior that we recognize as central to many species.

A first way epistemic imperatives become apparent is during the minimization of *variational free energy*. One of the ways to decompose free energy is to express it as a Gibbs energy expected under the approximate posterior minus the entropy of the approximate posterior. In other words, the agent is striving to *increase* entropy. While this seems paradoxical, the paradox disappears if one considers that this is the entropy of the agent's (approximate posterior) belief. This can be understood as the imperative to explain things as accurately as possible but also "keep options open" and avoid committing to any specific explanation unless this is necessary—that is, the *maximum entropy* principle (Jaynes 1957).

A second way epistemic dynamics become apparent is during the minimization of *expected free energy*, wherein—interestingly—there are two entropies with opposite signs. These include the posterior predictive entropy (how

Box 3.3 (continued)

uncertain I am about what outcomes I would encounter given a choice) that must be maximized—as for beliefs about states in the variational free energy—and the conditional entropy of outcomes given states (the ambiguity entailed by a policy) that must be minimized. While during the minimization of variational free energy the imperative is to *maximize* entropy of (present) beliefs, during the maximization of expected free energy the imperative is to select actions that *minimize* the ambiguity of (future) beliefs. This gives rise to epistemic, curious, novelty-seeking, and information-foraging behaviors, which support uncertainty resolution or improvement of the generative model—which in turn minimizes surprise in the long run (Seth 2013; Friston, Rigoli et al. 2015; Seth and Friston 2016; Schwartenbeck, Passecker et al. 2019).

sensory observations are generated. This model defines surprise—or better, a tractable proxy, variational free energy, which can be measured and minimized efficiently.

An Active Inference agent appears to perform (approximate) Bayesian inference under a generative model and to maximize evidence for its model—that is, it is a self-evidencing agent. The prospective bit of the inference is realized by selecting courses of actions or policies that are expected to minimize free energy in the future. This formalism leads to a novel view of (optimal) behavior in terms of the Hamiltonian principle of least Action—a (first) principle that connects Active Inference to the domains of statistical physics, thermodynamics, and nonequilibrium steady states.