2 The Low Road to Active Inference

My thinking is first and last and always for the sake of my doing.
—William James

2.1 Introduction

This chapter introduces Active Inference by starting from the Helmholtzian—
or perhaps Kantian—view of "perception as unconscious inference" (Helm-
holtz 1867) and related ideas that have emerged more recently under the
Bayesian brain hypothesis. It explains how Active Inference subsumes and
extends these ideas by treating not just perception but also action, planning,
and learning as problems of (Bayesian) inference and by deriving a principled
(variational) approximation to such otherwise intractable problems.

2.2 Perception as Inference

There is a long tradition of seeing the brain as a "predictive machine,"
or a statistical organ that infers and predicts external states of the world.
This idea dates back to the notion of "perception as unconscious infer-
ence" (Helmholtz 1866). More recently, this has been reformulated as the
"Bayesian brain" hypothesis (Doya 2007). From this perspective, percep-
tion is not a purely bottom-up transduction of sensory states (e.g., from the
retina) into internal representations of what is out there (e.g., as patterns
of neuronal activity). Rather, it is an inferential process that combines (top-
down) prior information about the most likely causes of sensations with
(bottom-up) sensory stimuli. Inferential processes operate on probabilistic

representations of states of the world and follow Bayes' rule, which pre-
scribes the (optimal) update in the light of sensory evidence. Perception is
not a passive outside-in process—in which information is extracted from
impressions on our sensory epithelia from "out there." It is a constructive
inside-out process—in which sensations are used to confirm or disconfirm
hypotheses about how they were *generated* (MacKay 1956, Gregory 1980,
Yuille and Kersten 2006, Neisser 2014, A. Clark 2015).

In turn, performing Bayesian inference requires a *generative model*—
sometimes referred to as a *forward model*. A generative model is a construct
from statistical theory that generates predictions about observations. It may
be formulated as the *joint probability* $P(y, x)$ of *observations y* and the world's
hidden states x that generate these observations. The latter are referred to
as *hidden* or *latent* states as they cannot be observed directly. This joint
probability can be decomposed into two parts. The first is a *prior* $P(x)$, which
denotes the organism's knowledge about the hidden states of the world prior
to seeing sensory data. The second is the *likelihood* $P(y|x)$, which denotes
the organism's knowledge of how observations are generated from states.
Bayes' rule tells us how to combine these two elements, essentially updating
a *prior probability* $P(x)$ into a *posterior probability* of hidden states after receiv-
ing observations $P(x|y)$. For readers who need a brief refresher on basic
probability theory, box 2.1 provides a summary.

Bayesian inference is a broad topic that arises in disciplines like statistics,
machine learning, and computational neuroscience. A full treatment of the
associated topics is beyond the scope of this book, but there are excellent
resources available for those who wish to understand it in depth (Murphy
2012). However, all of this is based on one simple rule. To illustrate this rule,
we consider an example of Bayesian perceptual inference (figure 2.1). Imag-
ine a person who has a strong belief that she is confronted with an apple.
This belief corresponds to a prior probability, or *prior* for short. This prior
comprises the probability attributed to the apple hypothesis and the prob-
ability assigned to alternative hypotheses. In this example, our alternative
hypothesis is that it is not an apple but a frog. Numerically, the prior prob-
ability distribution assigns 0.9 to apple and 0.1 to frog. Note that, as we have
assumed that there are only two plausible (mutually exclusive) hypotheses,
they must sum to one. The person is also equipped with a *likelihood* model,
which assigns a high probability to the fact that frogs jump, whereas apples
do not. This likelihood specifies the (probabilistic) mapping from the two
hidden states (frog or apple) to the two observations (jumps or does not

Box 2.1

The sum and product rules of probability

Probabilistic reasoning is underwritten by two key rules: the sum and product rules of probability, which are as follows (respectively):

$$\sum_x P(x) = 1$$
$$P(x)P(y|x) = P(x,y)$$

The sum rule says that the probability of all possible events (x) must sum (or integrate) to one. The product rule says that the *joint* probability of two random variables (x and y) may be decomposed into the product of the probability of one variable ($P(x)$) and the *conditional* probability of the second variable given the first ($P(y|x)$). A conditional probability is the probability of one variable (here, y) if we know the value that the other variable (here, x) takes.

We can develop two important results from these simple rules. The first is the operation of *marginalization*. The second is Bayes' rule. Marginalization allows us to obtain a distribution of just one of the two variables from a joint distribution:

$$\underbrace{\sum_x P(x,y) = \sum_x P(y)P(x|y)}_{\text{Product rule}} = \underbrace{P(y)\sum_x P(x|y) = P(y)}_{\text{Sum rule}}$$

The probability of y is referred to as a marginal probability, and we refer to this operation as marginalizing out x. Bayes' rule may be obtained directly from the product rule:

$$\underbrace{P(x)P(y|x)}_{\text{Product rule}} = P(x,y) = \underbrace{P(y)P(x|y)}_{\text{Product rule}}$$

This lets us translate between a prior and conditional distribution (likelihood) and the associated marginal and the other conditional distribution (posterior). Put simply, Bayes' rule just says that the probability of two things is the probability of the first, given the second, times the probability of the second, which is the same as the probability of the second, given the first, times the probability of the first.

jump). Together, the prior and the likelihood form the person's generative model.

Now imagine that the person observes that her apple-frog jumps. Bayes' rule tells us how to form a posterior belief from the prior, taking into account the likelihood of jumping. This rule is expressed as follows:

$$P(x|y) = \frac{P(x)P(y|x)}{P(y)} \tag{2.1}$$

Prior beliefs   **Posterior beliefs**

$P(x = frog)$ $P(x = apple)$ $P(x = frog \mid y = jumps)$ $P(x = apple \mid y = jumps)$

| 0.1 | 0.9 | Observe jumping | 0.9 | 0.1 |

Likelihood model

Apples   Frogs

$P(y = jumps \mid x = apple)$ 0.01 $P(y = jumps \mid x = frog)$

0.99   0.81

$P(y = doesn't\ jump \mid x = apple)$ 0.19 $P(y = doesn't\ jump \mid x = frog)$
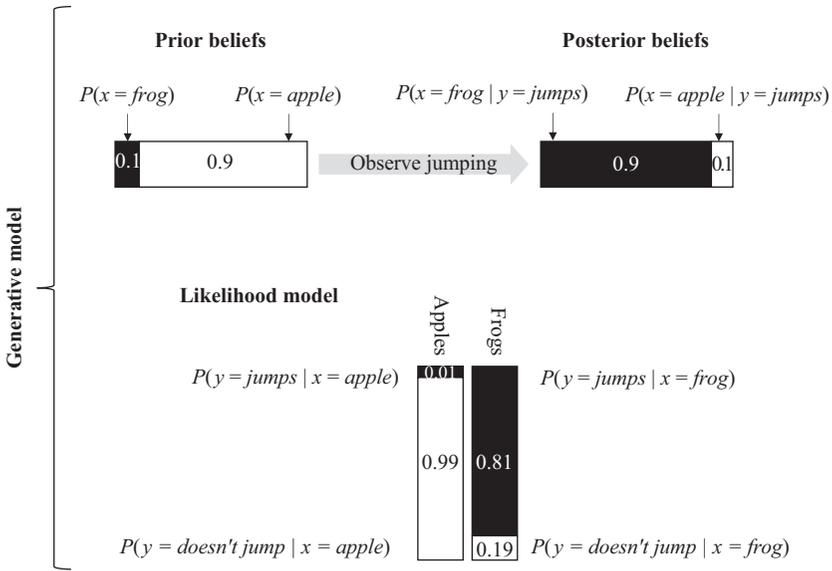
Generative model

Figure 2.1

A simple example of Bayesian inference. *Upper left:* The organism's prior belief $P(x)$ about the object it will see, before having made any observations, i.e., a categorical distribution over two possibilities, apple (with probability 0.9) and frog (with probability 0.1). *Upper right:* The organism's posterior belief $P(x \mid y)$ after observing that the object jumps. Posterior beliefs can be computed using Bayes' rule under a likelihood function $P(y \mid x)$. This is shown below the prior and posterior, and specifies that if the object is an apple, there is a very small probability (0.01) that it will jump, while if it is a frog, the probability that it will jump is much higher (0.81). (The probability bars in this figure are not exactly to scale.) In this specific case, the update from prior to posterior is large.

Under the likelihood model in figure 2.1, the posterior probability assigned to the frog is 0.9, and the probability assigned to the apple is 0.1. As highlighted in box 2.1, the denominator of equation 2.1 may be computed by marginalizing the numerator. Using our apple-frog example, we take the opportunity to unpack two different notions of *surprise*—both of which are important in Active Inference. The first, which we refer to simply as surprise, is the negative log evidence, where evidence is the marginal probability of observations. In our example, this is the negative log probability of observing anything jumping under the generative model. Surprise is a very important quantity from a Bayesian perspective. It is a measure of how poorly a model fits the data it tries to explain. To put this intuitively,

we can work out the probability of the observed (jumping) behavior under our model. Remember that this assigns a very high prior probability to apples and a low prior probability to frogs. Thus, our marginal probability of jumping is as follows:

$$
\begin{aligned}
P(y = jumps) &= \sum_x P(x, y = jumps) \\
&= \sum_x P(x)P(y = jumps \,|\, x) \\
&= P(x = frog)P(y = jumps \,|\, x = frog) \\
&\quad + P(x = apple)P(y = jumps \,|\, x = apple) \\
&= 0.1 \times 0.81 + 0.9 \times 0.01 \\
&= 0.09
\end{aligned}
\tag{2.2}
$$

This means that, under this model, we would only expect to observe jumping behavior about 9 times out of 100 observations. As such, we should be surprised to observe this if we subscribed to the model in figure 2.1. We can quantify this in terms of surprise ($\Im$). This is given by $\Im(y = jumps) = -\ln P(y = jumps) = -\ln(0.09) = 2.4$ nats.[1] The bigger this number, the worse the model as an apt explanation for the observations at hand. This lets us compare models in relation to data. For example, consider an alternative model, where we have a prior belief that frogs are seen 100 percent of the time. Following the same steps as in equation 2.2, we calculate a surprise of about 0.2 nats. This is a better model of these data, as the observation is much less surprising. The procedure of scoring models on the basis of their evidence (or surprise) is often referred to as Bayesian model comparison. For more complicated models, the form of the surprise may not be so simple. Table 2.1 provides the form of the surprise (omitting constants) for a range of probability distributions—in addition to the categorical probability in our example. Crucially, this lets us talk about surprise for probability distributions whose support[2] differs from the simple example used here. This is important because the way in which sensory data are generated by the world varies with the sort of data. We could be surprised by encountering the face of someone we did not expect to see (categorical distribution), or we could be surprised by it being colder outside than we anticipated (continuous distribution). Table 2.1 may be seen as a portfolio of the probability distributions at our disposal when we come to construct generative models in subsequent chapters. More generally, it makes the point that surprise is a concept that can be evaluated for any given family of probability distributions.

Table 2.1
Probability distributions and surprise[3]

Distribution	Support	Surprise ($\mathfrak{I}$)
Gaussian	$x \in \mathbb{R}$	$\frac{1}{2}(x - \mu) \cdot \Pi(x - \mu)$
Multinomial[1]	$x_i \in (0, \dots, N)$ $i \in \{1, \dots, K\}$ $\sum_i x_i = N$	$-\sum_i x_i \ln d_i$
Dirichlet[2]	$x_i \in (0,1)$ $i \in \{1, \dots, K\}$ $\sum_i x_i = 1$	$\sum_i (1 - \alpha_i) \ln x_i$
Gamma	$x \in (0, \infty)$	$(bx + (1-a) \ln x)$

Notes: 1. Special cases include categorical ($K > 2$, $N = 1$), binomial ($K = 2$, $N > 1$), and Bernoulli ($K = 2$, $N = 1$) distributions. 2. A special case is the beta distribution ($K = 2$).

The second notion of surprise is (slightly confusingly) referred to as *Bayesian surprise*. This is a measure of how much we have to update our beliefs following an observation. In other words, Bayesian surprise quantifies the difference between a prior and a posterior probability. This raises the question of how we quantify the dissimilarity of two probability distributions. One answer, from information theory, is to use a Kullback-Leibler (KL) Divergence. This is defined as the average difference between two log probabilities:

$$D_{KL}\big[Q(x) \| P(x)\big] \triangleq \mathbb{E}_{Q(x)}\big[\ln Q(x) - \ln P(x)\big] \tag{2.3}$$

The $\mathbb{E}$ symbol here indicates an average (or expectation) as outlined in box 2.2. Using the KL-Divergence, we can quantify the Bayesian surprise of our example:

$$\begin{aligned}
&D_{KL}\big[P(x|y) \| P(x)\big] \\
&= P(x = frog | y = jumps)\big(\ln P(x = frog | y = jumps) - \ln P(x = frog)\big) \\
&\quad + P(x = apple | y = jumps)\big(\ln P(x = apple | y = jumps) - \ln P(x = apple)\big) \\
&= 0.9\big(\ln(0.9) - \ln(0.1)\big) + 0.1\big(\ln(0.1) - \ln(0.9)\big) \\
&\approx 1.8 \text{ nats}
\end{aligned} \tag{2.4}$$

This scores the amount of belief updating, as opposed to simply how unlikely the observation was. To highlight the distinction between surprise and Bayesian surprise, consider what happens if we commit to a prior belief that we will always see apples. The Bayesian surprise will be zero, as the prior

Box 2.2
Expectations

It is useful to refer to the *expectation* of a random variable x, usually denoted $\mathbb{E}[x]$. This is the weighted average of all the values that the variable can assume, weighted by their probability. For discrete random variables (that can only take a countable number of possible values), this is given by a weighted sum:

$$\mathbb{E}[x] = \sum_x xP(x)$$

For example, for a discrete (numerical) variable that can only assume two values (1 and 2) with equal probability of $\frac{1}{2}$, this is $\mathbb{E}[x] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2}$.

For continuous random variables (that can take infinitely many values), sums are replaced by integrals. Expectations can also be applied to functions of random variables, as opposed to the variables directly. For example, if we have a function $f(x)$, where x has some continuous distribution, the expectation is defined as follows:

$$\mathbb{E}[f(x)] = \int f(x)p(x)\,dx$$

We will use this notation throughout this book, where the function $f(x)$ will often be a log probability, or log probability ratio.

is so confident that we do not update it at all following our observations. However, the surprise is very large (4.6 nats) as it is highly unlikely that an apple will jump.

Note that while we illustrated Bayesian inference on the basis of a very simple generative model, it applies to generative models of any complexity. In chapter 4 we will highlight two forms of generative model that underwrite most applications in Active Inference.

2.3 Biological Inference and Optimality

There are two important points that connect the above inferential scheme to biological and psychological theories of perception. First, the inferential procedure discussed requires the interplay of top-down processes that encode predictions (from the prior) and bottom-up processes that encode sensory observations (as mediated by the likelihood). This interplay of top-down and bottom-up processes distinguishes the inferential view from alternative approaches that only consider bottom-up processes. Furthermore, it

is central in modern biological treatments of perception, such as *predictive coding* (discussed in chapter 4), which is a specific algorithmic (or process-level) implementation of the more general (Bayesian) inference scheme discussed here.

Second, Bayesian inference is *optimal*. Optimality is defined in relation to a cost function that is optimized (i.e., minimized), which, for Bayesian inference, is known as variational free energy—closely related to surprise. We return to this in section 2.5. By explicitly considering the full distribution over hidden states, it naturally handles uncertainty, hence avoiding the limitations of alternative approaches that only consider point estimates of hidden states (e.g., the mean value of *x*). One such alternative would be maximum likelihood estimation, which simply selects the hidden state most likely to have generated the data at hand. The problem with this is that such estimates ignore both the prior plausibility of the hidden state and the uncertainty surrounding the estimation. Bayesian inference does not suffer these limitations. However, despite the use of surprise in objectively assessing whether the model is fit for purpose, it is important to appreciate that inference itself is *subjective*. The results of inference are not necessarily accurate in any objective sense (i.e., the organism's belief may not actually correspond to reality) for at least two important reasons. First, biological creatures operate on the basis of limited computational and energetic resources, which render exact Bayesian inference intractable.[4] This requires approximations that preclude guarantees of exact Bayesian optimality. These approximations include the notion of a variational posterior—based on something called a mean field approximation—which is central to chapter 4.

The second reason optimality may be thought of as subjective is that organisms operate on the basis of a subject's *generative model* of how their observations are generated, which may or may not correspond to the real *generative process* that generates their observations. This is not to say that the *generative model* should correspond to the *generative process*. In fact, there may be models that afford better (e.g., simpler) explanations of the data at hand than the processes that actually generated them—as quantified by their relative surprise. A nice example of this is illusions, for which someone finds a simpler explanation for their visual input in relation to how the visual stimuli have been carefully engineered by a mischievous psychophysicist.

The generative model itself may be optimized as new experience is acquired. This may or may not converge to the generative process. Figure 2.2 illustrates
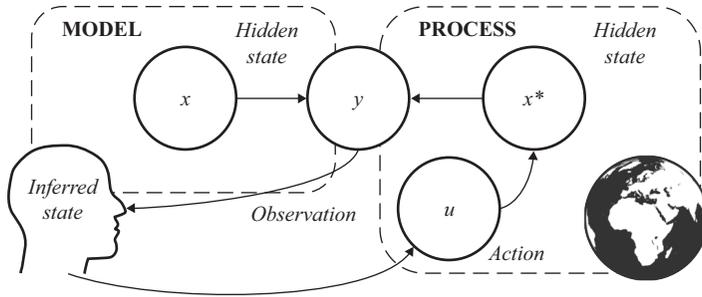
Figure 2.2

Generative process and generative model. Both represent ways in which sensory data (y) could be generated given hidden states (x) and are represented through arrows from x to y to indicate causality. The difference is that the process is the true causal structure by which data are generated, while the model is a construct used to draw inferences about the causes of data (i.e., use observations to derive inferred states). The hidden states of the generative model and the generative process are not the same. The organism's model includes a range of hypotheses (x) about the hidden state, which do not necessarily include the true value of the hidden state x^* of the generative process. In other words, the models we use to explain our sensorium may include hidden states that do not exist in the outside world, and vice versa. Action (u) is generated on the basis of the inferences made under a generative model. Action is shown here as part of the generative process, making changes to the world, despite being selected from the inferences drawn under the model.

this point and the difference between true environmental contingencies, or the generative process, which is inaccessible to the organism and the organism's generative model of the world. In this particular example, the generative process is in a true state x^* that is inaccessible to the organism. However, the organism and world are reciprocally coupled, and x^* generates an observation y, which the organism senses. The organism can use this observation y and Bayes' rule to infer the (posterior probability of) some explanatory variable or hidden state in the generative model. In the figure, we refer to both x^* and x as *hidden states*, emphasizing that neither is observable. However, they are subtly different: the former is part of the organism's generative model, whereas the latter is part of the generative process and inaccessible to the organism. Furthermore, x^* and x do not necessarily live in the same space. It might be that the hidden states in the external world take on values that lie outside the space of explanations available to the brain. Conversely, it might be that the brain's explanations include variables that do not exist in the

outside world. For example, the former could be 5-dimensional and the latter 2-dimensional, or one could be continuous and the other categorical.

The distinction between the generative model and process is important to contextualize psychological claims about optimality of inference—to the extent that these claims are valid—which, on a Bayesian view, is always contingent on the organism's resources. By resources, we mean its specific generative model, and bounded computational and mnemonic resources.

2.4   Action as Inference

The discussion to this point is common to all Bayesian brain theories. However, we now introduce the simple but fundamental advance offered by Active Inference. This starts from the same inferential perspective discussed above but extends it to consider action as inference. This idea stems from the concept that Bayesian inference minimizes surprise (or, equivalently, maximizes Bayesian model evidence). So far, we have considered what happens when we compute surprise by performing inference—and select among models on the basis of their capacity to minimize surprise. However, surprise does not only depend on the model. It also depends on the data. By acting on the world to change the way in which data are generated, we can ensure a model is fit for purpose by choosing those data that are least surprising under our model.

Equipped with a mechanism to produce actions, an organism can engage in reciprocal exchanges with its environment; see figure 2.2. In animals, this mechanism takes the form of a motor reflex loop. Essentially, for each action-perception cycle, the environment sends an observation to the organism. The organism uses (an approximation to) Bayesian inference to infer its most likely hidden states. It then generates an action and sends it to the environment in an attempt to make the environment less surprising. The environment executes the action, generates a new observation, and sends it to the organism. Then, a new cycle starts. The sequential description here is written for didactic purposes; it is important to realize that these are not really discrete steps but are continuous dynamical processes.

Active Inference goes beyond the recognition that perception and action have the same (inferential) nature. It also assumes that both *perception and action cooperate to realize a single objective*—or optimize just one function—rather than having two distinct objectives, as more commonly assumed. In

the Active Inference literature, this common objective has been described in various (informal and formal) ways, including the minimization of surprise, entropy, uncertainty, prediction error, or (variational) free energy. These terms are related to one another but sometimes their relations are not immediately clear, causing some confusion. Furthermore, these terms are used in different contexts; for example, *prediction error* minimization is used in biological contexts where the objective is explaining brain signals, while variational free energy minimization is used in machine learning.

In the next two sections, we will clarify that the single quantity that Active Inference agents minimize through perception and action is *variational free energy*. However, under some conditions, one can reduce variational free energy to other notions, such as the discrepancy between the generative model and the world, or the difference between what one expects and what one observes (i.e., a prediction error). We will introduce variational free energy formally in section 2.5. For simplicity section 2.4 focuses on the ways in which perception and action minimize the discrepancy between the generative model and the world.

2.5   Minimizing the Discrepancy between Model and World

Having established perception and action in terms of Bayesian inference, we now turn to the question of what the objective of inference is. In other words, what is being optimized by inference? In cognitive science, it is common to assume that different cognitive functions like perception and action optimize different objectives. For example, we could assume perception maximizes the accuracy of reconstruction while action selection maximizes utility. Instead, a fundamental insight of Active Inference is that both perception and action serve the very same objective. As a first approximation, this common objective of *perception and action* can be formulated as a minimization of the *discrepancy between the model and the world*. Sometimes this is operationalized in terms of prediction error.

To understand how perception and action reduce the discrepancy between the model and the world, consider again the example of a person who expects to see an apple (figure 2.3). She generates a top-down visual prediction (e.g., about seeing something red and not jumping). This visual prediction is compared with a sensation (e.g., something jumping)—and this comparison results in a discrepancy.
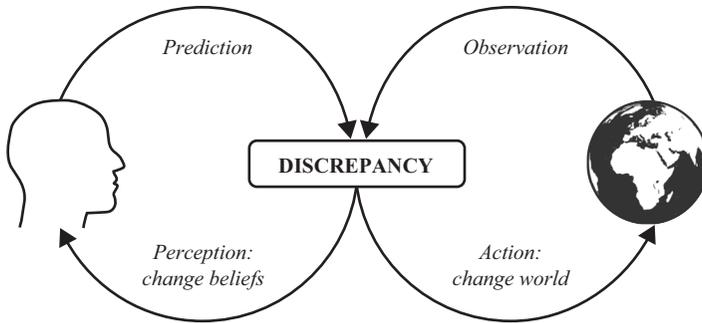
Figure 2.3
Both perception and action minimize discrepancy between model and world.

The person can resolve this discrepancy in two ways. First, she could change her mind about what she is seeing (i.e., a frog) to fit the world, hence resolving the discrepancy. This corresponds to *perception*. Second, she could foveate the nearest apple tree and see something that looks very much like an apple. This also resolves the initial discrepancy, but in a different way. This entails changing the world—including her direction of gaze— and subsequent sensations to fit what is in her mind, not changing her mind to fit the world. This is the other direction of fit. This is *action*.

While changing the direction of one's gaze seems less compelling than changing one's mind in the world of apples and frogs, let us consider another case: a person who expects his body temperature to be in a certain range who senses a high temperature via central thermoreceptors. This is surprising and presents a significant discrepancy to resolve. As in the former example, he has two ways to minimize this discrepancy, corresponding to perception (changing mind) and action (changing the world), respectively. In this case, simply changing one's mind does not seem very adaptive, but acting to lower the body temperature (e.g., by opening the window) is.

This speaks to the fact that in Active Inference, the notion of marginal probabilities or *surprise* (e.g., about body temperature) has a meaning that goes beyond standard Bayesian treatments to absorb notions like homeostatic and allostatic set-points. Technically, Active Inference agents come equipped with models that assign high marginal probabilities to the states they prefer to visit or the observations they prefer to obtain. For a fish, this means a high marginal likelihood for being in water. This implies that

organisms implicitly expect the observations they sample to be within their comfort zone (e.g., physiological bounds).

In sum, we have discussed how, at any point in time, we can minimize the discrepancy between our model and our world through perception and action. Whether we adjust our beliefs or our data depends on the confidence with which we hold those beliefs. In our example of the apple, the belief is held with sufficient uncertainty that this will be updated as opposed to acted on. In contrast, in the temperature example, we are considerably more confident about our core temperature because it underwrites our existence. This confidence means we update our world to comply with our beliefs. Yet, in Active Inference, perception and action act more cooperatively than suggested by this treatment. To understand why this is the case, the next section moves from the restricted notion of *discrepancy* (or *prediction error*) to the more general notion of *variational free energy*—which is the quantity that Active Inference actually minimizes and which subsumes prediction error as a special case.

2.6   Minimizing Variational Free Energy

So far, we have discussed perception and action within a Bayesian scheme that aims to minimize surprise. Yet, exact Bayesian inference supporting perception and action is computationally intractable in most cases, because two quantities—the *model evidence* ($P(y)$) and the *posterior probability* ($P(x|y)$)—cannot be computed for two possible reasons. The first is that for complex models, there may be many types of hidden states that all need marginalizing out, making the problem computationally intractable. The second is that the marginalization operation might require analytically intractable integrals. Active Inference appeals to a *variational* approximation of Bayesian inference that is tractable.

The formalism of variational inference will be unpacked in chapter 4. Here, it suffices to say that performing variational Bayesian inference implies substituting the two intractable quantities—*posterior probability* and (log) *model evidence*—with two quantities that approximate them but can be computed efficiently—namely, an *approximate posterior Q* and a *variational free energy F*, respectively. The approximate posterior is sometimes called a variational or recognition distribution. Negative variational free energy

is also known as an evidence lower bound (ELBO), especially in machine learning.

Most importantly, the problem of Bayesian inference now becomes a problem of optimization: the minimization of *variational free energy F*. Variational free energy is a quantity with roots in statistical physics that plays a fundamental role in Active Inference. In equation 2.5, it is denoted as $F[Q, y]$, as it is a functional (function of a function) of the approximate posterior Q and a function of data y:

$$
\begin{aligned}
F[Q, y] &= \underbrace{-\mathbb{E}_{Q(x)}[\ln P(y, x)]}_{Energy} - \underbrace{H[Q(x)]}_{Entropy} \\
&= \underbrace{D_{KL}[Q(x) \,\|\, P(x)]}_{Complexity} - \underbrace{\mathbb{E}_{Q(x)}[\ln P(y \,|\, x)]}_{Accuracy} \\
&= \underbrace{D_{KL}[Q(x) \,\|\, P(x \,|\, y)]}_{Divergence} - \underbrace{\ln P(y)}_{Evidence}
\end{aligned}
\tag{2.5}
$$

Variational free energy may seem prima facie an abstract concept, but its nature and the role it plays in Active Inference become apparent when decomposed into quantities that are more intuitive and familiar in cognitive science. Each of these perspectives on variational free energy offers useful intuitions about what free energy minimization means. We briefly sketch these intuitions here, as they will become important when we discuss examples in the second part of the book.

The first line of equation 2.5 shows that minimizing with respect to Q requires consistency with the generative model (*energy*) while also maintaining a high posterior *entropy*.[5] The latter means that, in the absence of data or precise prior beliefs (which only influence the *energy* term), we should adopt maximally uncertain beliefs about the hidden states of the world, in accord with Jaynes's *maximum entropy* principle (Jaynes 1957). Put simply, we should be uncertain (adopt a high entropy belief) when we have no information. The term *energy* inherits from statistical physics. Specifically, under a Boltzmann distribution, the average log probability of a system adopting some configuration is inversely proportional to the energy associated with that configuration—that is, the energy required to move the system into this configuration from a baseline configuration.

The second line emphasizes the interpretation of free energy minimization as finding the best explanation for sensory data, which must be the simplest (minimally *complex*[6]) explanation that is able to *accurately*[7] account for the data (cf. Occam's razor). The complexity-accuracy trade-off

recurs across several domains, normally in the context of model comparison for data analysis. In statistics, other approximations to model evidence are sometimes used, such as the Bayesian information criterion or Akaike information criterion. The complexity-accuracy trade-off will become important when we describe how to use free energy for model comparison during model-based data analysis—and in the context of structure learning and model reduction. Inferring explanations that have minimal complexity is also important from a cognitive perspective. This is because one can assume that updating what one knows (the prior) to accommodate the data entails a cognitive cost (Ortega and Braun 2013, Zénon et al. 2019); hence, an explanation that diverges minimally from the prior is preferable.

On this view, the complexity cost is just Bayesian surprise. In other words, the degree to which "I change my mind" is quantified by the divergence between the prior and the posterior. This means every accurate explanation for my sensations incurs a complexity cost, and this cost scores the degree of *Bayesian belief updating*. Variational free energy, then, scores the difference between accuracy and complexity.

The final line expresses the free energy as a bound on negative log evidence (see figure 2.4). As the left part of the figure illustrates, the free energy is an upper bound on negative log evidence, where the bound is the divergence between Q and the posterior probability that would have been obtained were it possible to perform exact (as opposed to variational) inference. The right part of the figure shows that as the divergence decreases, the
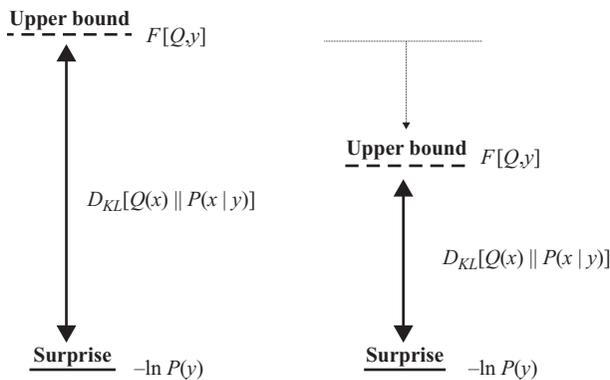

Figure 2.4
Variational free energy as an upper bound on negative log evidence.

free energy approaches the negative log evidence (surprise)—and becomes equal to surprise, if the approximate posterior Q matches the exact posterior $P(x|y)$. This offers a formal motivation for perceptual inference as one way to lower free energy by optimizing our approximate posterior Q as much as possible.

The final line of equation 2.5 shows that perceptual inference is not the only way to minimize free energy. We could also change the log evidence term through acting to change sensory data. This decomposition is interesting from a cognitive perspective, since *minimizing divergence* and *maximizing evidence* map to the two complementary sub-objectives of perception and action, respectively; see figure 2.5. Note that the above expressions all become ways of characterizing the negative log evidence if we replace Q with $P(x|y)$, generalizing to the case of exact inference.

In sum, Active Inference amounts to minimizing variational free energy by perception and action. This minimization permits an organism to fit its generative model to the observations it samples. This fit is a measure of both perceptual adequacy (as expressed by the divergence term) and active control over external states—in the sense that it permits the organism to maintain itself in a suitable set of preferred states, as defined by the generative model. Another way of phrasing this is to appeal to the divergence versus
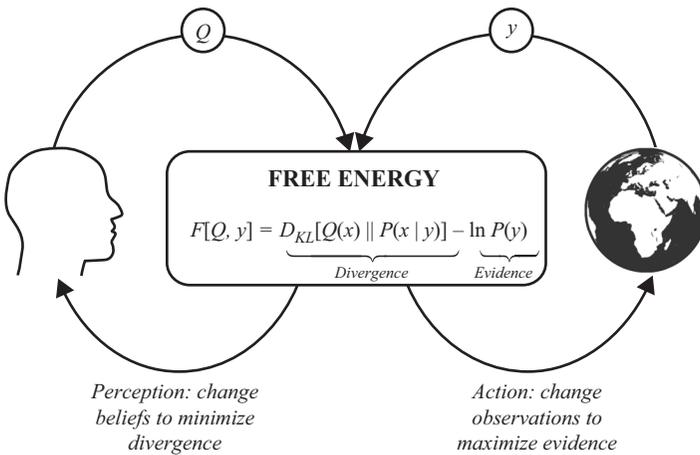

Figure 2.5
Complementary roles of perception and action in the minimization of variational free energy.

evidence decomposition of free energy. Equating the negative log evidence with surprise, and noting that the smallest possible divergence is zero, we see that free energy is an upper bound on surprise. This means it can only be greater than or equal to surprise. When the organism minimizes its divergence (through perception), then free energy becomes an approximation to surprise. When an organism additionally changes the observations it gathers (by acting) to render them more similar to prior predictions, it minimizes surprise.

Variational free energy has a *retrospective* aspect, as it is a function of past and present, but not future, observations. Although it facilitates inferences about the future based on past data, it does not directly facilitate *prospective* forms of inference based on anticipated future data. This is important in planning and decision-making. Here, we infer the best actions or action sequences (policies) on the basis of the future observations they are expected to bring about. Doing this requires that we supplement our generative models with the notion of *expected free energy*.

2.7 Expected Free Energy and Planning as Inference

Expected free energy extends Active Inference to include a quintessentially prospective form of cognition: planning. Planning a sequence of actions, such as the series of moves required to escape from a maze, requires considering future observations that one expects to gather. For example, the consequences of possible courses of action include seeing a dead end after turning right or seeing the exit after a sequence of three left turns. Each possible sequence of actions is termed a *policy*. This highlights an important distinction made in Active Inference between an action and a policy. The former is something that directly influences the outside world, while the latter is a hypothesis about a way of behaving. The implication is that Active Inference treats planning and decision-making as a process of inferring what to do. This brings planning firmly into the realm of Bayesian inference and means we must specify priors and likelihoods as before (section 2.1). However, in place of frogs and apples, the alternatives are behavioral policies (Is it more probable that I look toward the pond or the tree?). In this section, we first briefly deal with the likelihood—that is, the consequences of pursuing a policy—and then turn to the prior. This is where expected free energy comes in.

Policy-dependent outcomes are not immediately available (they are in the future), but they can be predicted by chaining together two components of the generative model. The first is our beliefs about how hidden states change as a function of policies. We will get into the details of this in chapter 4. For now, we use the notation $\tilde{x}$ to denote a sequence or trajectory of hidden states over time, and we condition trajectories on the policies (π) a creature pursues. This means the dynamical part of our model is given by $P(\tilde{x}|\pi)$. Drawing from our earlier frog-apple example, the policy may be the decision to go to a pond or to an orchard, which changes the probability of encountering frogs versus apples.

The second component of the model is the usual *likelihood* distribution. This describes which observations to expect in every possible state (e.g., jumping or not, conditioned on frog or apple). By combining these two components, an organism can engage its generative model vicariously to run "what if" or *counterfactual* simulations of the consequences of its possible actions or policies—for example, "What would happen if I go to the pond?" Marginalizing over states, this gives us the marginal likelihood or evidence for a policy ($P(\tilde{y}|\pi)$), or a free energy approximation to this quantity. In other words, knowing how policies influence state transitions lets us compute the likelihood of a sequence of observations under that policy. As we saw in equation 2.1, we need to combine this likelihood with a prior probability to calculate the posterior probability of pursuing a policy.

Active Inference decomposes this planning problem into two successive operations. The first is to compute a score for each policy. The second is to form posterior beliefs about which to pursue. The former defines the prior belief about the policies to pursue, where the best policies have high probability and the worst policies have low probability. Under Active Inference, the goodness of a policy is scored by the associated negative expected free energy—just as the goodness of a model fit is scored by the negative free energy of that model. The *expected free energy* (G) of policy is different from the variational free energy (F), since calculating the former requires consideration of future, policy-dependent observations. In contrast, the latter only considers present and past observations. Calculating expected free energy therefore engages the generative model to predict future observations that would stem from each policy—if it were to be executed—up to some planning horizon. Furthermore, because a policy unfolds over multiple time

steps, the final measure of expected free energy for each policy has to integrate over all future time steps of that policy.

The expected free energy of each policy can be converted in a quality score (by taking its negative) and is directly available as a prior by agents engaging in Active Inference. This is because—consistent with the notion of potential energy in physics—expected free energy is expressed in the space of log probabilities. Converting it into a belief (or probability distribution) over policies is then a matter of exponentiating (to undo the log) and normalizing (to ensure consistency with the sum rule in box 2.1). Policies that are associated with a lower expected free energy are assigned higher probability and become the policies that the organism expects to pursue.

Ultimately, inferring that we are pursuing a particular policy has consequences for the sensory data we predict. For example, a policy that includes flexing my elbow entails predictions about the proprioceptive input from the biceps and triceps muscles. This provides the link between planning and action, as the predictions associated with a plan translate into action that resolves discrepancies with measured proprioceptive data (see section 2.3).

2.8   What Is Expected Free Energy?

So far, we have assumed that during planning, the organism scores its policies according to their expected free energy. However, we have sidestepped what expected free energy actually is. Like *variational free energy*, the expected free energy can be decomposed in several, mathematically equivalent ways. Each of these provides an alternative perspective on this quantity.

$$
\begin{aligned}
G(\pi) = & -\underbrace{\mathbb{E}_{Q(\tilde{x},\tilde{y}|\pi)}[D_{KL}[Q(\tilde{x}\,|\,\tilde{y},\pi)\,\|\,Q(\tilde{x}|\pi)]]}_{\text{Information gain}} - \underbrace{\mathbb{E}_{Q(\tilde{y}|\pi)}[\ln P(\tilde{y}|C)]}_{\text{Pragmatic value}} \\
= & \underbrace{\mathbb{E}_{Q(\tilde{x}|\pi)}[H[P(\tilde{y}|\tilde{x})]]}_{\text{Expected ambiguity}} + \underbrace{D_{KL}[Q(\tilde{y}|\pi)\,\|\,P(\tilde{y}|C)]}_{\text{Risk (outcomes)}} \\
\leq & \underbrace{\mathbb{E}_{Q(\tilde{x}|\pi)}[H[P(\tilde{y}|\tilde{x})]]}_{\text{Expected ambiguity}} + \underbrace{D_{KL}[Q(\tilde{x}|\pi)\,\|\,P(\tilde{x}|C)]}_{\text{Risk (states)}} \\
= & -\underbrace{\mathbb{E}_{Q(\tilde{x},\tilde{y}|\pi)}[\ln P(\tilde{y},\tilde{x}|C)]}_{\text{Expected energy}} - \underbrace{H[Q(\tilde{x}|\pi)]}_{\text{Entropy}}
\end{aligned}
\tag{2.6}
$$

$$Q(\tilde{x},\tilde{y}\,|\,\pi) \triangleq Q(\tilde{x}|\pi)P(\tilde{y}|\tilde{x})$$

The first of these is perhaps the most useful, intuitively, as it expresses the value of seeking new information (i.e., exploration) in exactly the same

units (nats) as the value of seeking preferred observations (i.e., exploitation), dissolving the classic exploit-explore dilemma in behavioral psychology. By minimizing expected free energy, the relative balance between these terms determines whether behavior is predominantly explorative or exploitative. Note that pragmatic value emerges as a prior belief about observations, where the *C*-parameter includes preferences. The (potentially unintuitive) link between prior beliefs and preferences is unpacked in chapter 7; for now, we note that this term can be treated as an expected utility or value, under the assumption that valuable outcomes are the kinds of outcomes that characterize each agent (e.g., a body temperature of 37 °C).

The information gain term inherits from the divergence we considered in section 2.5, which ensures that free energy is an upper bound on surprise. However, there is a twist: instead of minimizing the divergence, we want to select policies that maximize the expected divergence—hence, information gain. This switch is due to the fact that we are now taking an average of the log probabilities over outcomes that have yet to be observed. This is a subtle point that can be understood in terms of outcomes switching their roles. When evaluating the free energy of outcomes, the outcomes are the consequences. However, when evaluating the expected free energy, the outcomes play the role of causes in the sense they are variables that are hidden in the future but explain decisions in the present.

The ensuing information gain penalizes observations for which there is a many-to-one mapping from observations to states—in the sense that one can obtain the same observations in different states—as this precludes precise belief updating. In artificial intelligence and robotics, states that bring the same observation (e.g., two T-junctions of a maze that look identical) are sometimes called *aliased* and are generally hard to deal with using simple methods (i.e., stimulus-response, with no inference or memory). The problem is that we cannot know which state we occupy from current observations alone. Active Inference avoids getting into such situations in the first place, given their low potential for information gain.

A simple example may help unpack the distinction between information gain (or epistemic value) and pragmatic value and highlight why, in most realistic situations, pragmatic and epistemic values need to be pursued in tandem. Imagine a person who wants an espresso and knows that there are two good cafes in the town: one that opens only from Monday to Friday and another that opens only during the weekend. If he does not know what day

of the week it is, he has to first select an action that has epistemic value and resolves his uncertainty (i.e., an *epistemic action* to look at the calendar)—and only after that select an action that carries pragmatic value and brings the reward (i.e., a *pragmatic action* to go to the correct cafe). This scenario illustrates the fact that in most uncertain situations, one must first perform epistemic actions to resolve uncertainty before confidently selecting a pragmatic action. Policy selection methods that fail to consider the epistemic affordance of choices can only select policies by using random number generators—and will often miss out on their espresso. Therefore, schemes that consider only pragmatic value are generally restricted to situations with no epistemic uncertainty, such as in the case of a person who already knows the day of the week and hence can head directly to the correct cafe.

The second decomposition in equation 2.6 is in terms of *risk* and expected *ambiguity*. These terms are the analogues of *complexity* and *inaccuracy*: risk is the expected complexity, and ambiguity is the expected inaccuracy. Risk, a common notion in economics, corresponds to the fact that there can be a one-to-many mapping between policies and their consequences—in the sense that one can obtain several different outcomes (by chance) under the same policy. One example is a gambling scenario with stochastic rewards (e.g., a one-armed bandit, aka a slot machine), wherein one could know the reward distribution—say, that one will obtain reward 10 percent of the time. This is called a risky situation in economics because, after the same move (pulling a lever), one could obtain two different observations (reward or no reward). This means one has to choose policies or plans that accommodate uncertainty. In risk-sensitive schemes—like active inference—the game is to choose policies whose probabilistic outcomes match, in the sense of a KL-Divergence, one's prior preferences. In short, minimizing complexity cost becomes minimizing risk when both are measures of departure from prior beliefs.

Similarly, *ambiguity* corresponds to the expected inaccuracy due to an ambiguous mapping between states and outcomes. A mapping is ambiguous if the distribution of outcomes anticipated is highly dispersed (or entropic) even if we know the states generating them with complete confidence. For instance, the probability of heads or tails in a coin flip, conditioned on whether it is sunny or raining, will be maximally ambiguous as there is no relationship between the weather and the 50-50 chance of heads or tails. As such, it would not be possible to gain information about the weather

by observing tails. Note that most situations are endowed with both risk and ambiguity—which implies a many-to-one mapping between states and outcomes and between policies and outcomes. Recall that outcomes (observations) are the only sort of variable that can be observed. Active Inference deals automatically with these situations, because expected free energy comprises both risk and ambiguity terms.

The third line of equation 2.6 highlights an alternative formulation of the expected free energy by reexpressing risk as a divergence between beliefs about states and preferences defined in terms of states. An appealing feature of this form is that it may be rearranged into an expected energy and entropy in analogy with variational free energy (equation 2.5). While this relationship is attractive, a downside of this formulation is that it assumes the state-space is known a priori such that prior preferences may be associated with states. In most settings, this is not a problem, and the choice between defining preferences in terms of states or outcomes has little practical relevance. However, common practice is to specify preferences in terms of outcomes—allowing the state-space itself to be learned while preserving extrinsic motivation.

In summary, expected free energy can be decomposed in terms of risk and ambiguity and in terms of pragmatic and epistemic values. These decompositions are interesting as they permit a formal understanding of the wide variety of situations that Active Inference deals with. Furthermore, they facilitate an appreciation of how Active Inference subsumes several decision schemes—which may be obtained by ignoring one or more components of expected free energy (figure 2.6). If one removes prior preferences, the pragmatic value becomes irrelevant, and all action is motivated by epistemic affordances—hence such schemes can only handle the resolution of uncertainty. Once prior preferences are removed, the (negative) expected free energy is variously known as expected *Bayesian surprise* (in the context of attentional exploration) or *intrinsic motivation* (in the context of autonomous learning). If one removes ambiguity, the resulting scheme corresponds to *risk-sensitive* or *KL control* in control theory. Finally, if one removes both ambiguity and prior preferences, the only remaining imperative is to maximize the entropy of observations (or states, if using the formulation in the third line of equation 2.6). This may be interpreted as uncertainty sampling (or keeping one's options open). Active Inference evinces the formal relations between these schemes and the (limited) situations in which they apply.
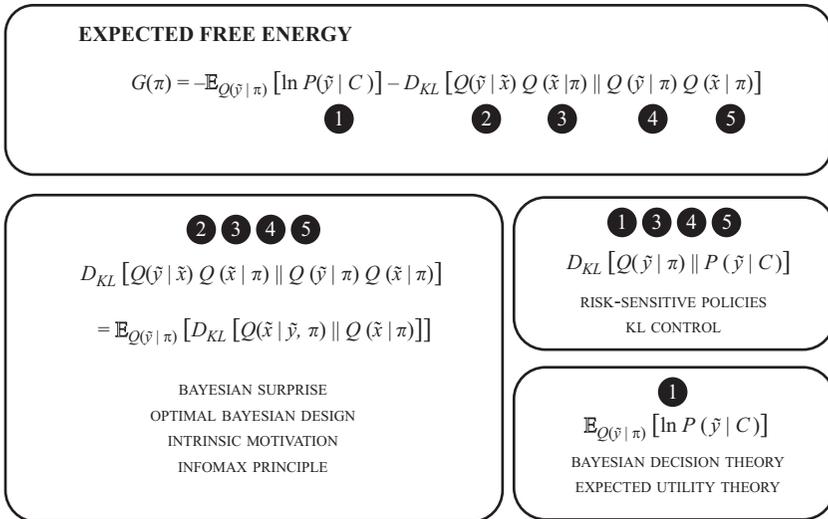
EXPECTED FREE ENERGY

$$G(\pi) = -\mathbb{E}_{Q(\tilde{y} \mid \pi)} \left[\ln P(\tilde{y} \mid C) \right] - D_{KL} \left[Q(\tilde{y} \mid \tilde{x}) \, Q\,(\tilde{x} \mid \pi) \parallel Q\,(\tilde{y} \mid \pi) \, Q\,(\tilde{x} \mid \pi) \right]$$

①   ②   ③   ④   ⑤

②③④⑤

$$D_{KL} \left[Q(\tilde{y} \mid \tilde{x}) \, Q\,(\tilde{x} \mid \pi) \parallel Q\,(\tilde{y} \mid \pi) \, Q\,(\tilde{x} \mid \pi) \right]$$

$$= \mathbb{E}_{Q(\tilde{y} \mid \pi)} \left[D_{KL} \left[Q(\tilde{x} \mid \tilde{y},\, \pi) \parallel Q\,(\tilde{x} \mid \pi) \right] \right]$$

BAYESIAN SURPRISE
OPTIMAL BAYESIAN DESIGN
INTRINSIC MOTIVATION
INFOMAX PRINCIPLE

①③④⑤

$$D_{KL} \left[Q(\tilde{y} \mid \pi) \parallel P\,(\tilde{y} \mid C) \right]$$

RISK-SENSITIVE POLICIES
KL CONTROL

①

$$\mathbb{E}_{Q(\tilde{y} \mid \pi)} \left[\ln P\,(\tilde{y} \mid C) \right]$$

BAYESIAN DECISION THEORY
EXPECTED UTILITY THEORY

Figure 2.6

Various schemes that can be derived by removing terms from the free energy equation. The upper panel shows the terms contributing to the expected free energy. The lower panels show the schemes that result from removing prior preferences (1), ambiguity (2), or everything except for the prior preferences. Each of these quantities appears in several different fields under a variety of names, but all can be seen as components of the same expected free energy.

Although we have carefully decomposed expected free energy in a way that different people might read this functional, there is no right or wrong way of carving it up. We will see in the second half of this book why autonomous systems of a certain kind must, in virtue of existing, choose actions that look as if they are minimizing expected free energy. This perspective means there is no privileged role for epistemic (explorative) versus pragmatic (exploitative) imperatives—or for risk versus ambiguity. These (possibly false) dichotomies are just two sides of the same existential coin.

2.9   At the End of the Low Road

Having introduced the two distinct notions of *variational free energy* and *expected free energy*, we are now in a position to consider what they achieve together. This represents an endpoint to the low road into Active Inference, starting from the notion of unconscious inference, via the Bayesian brain, the duality of perception and action, and finally planning as inference.

Variational free energy is at the core of Active Inference. It measures the fit between the internal generative model and (current and past) observations. By minimizing variational free energy, creatures maximize their model evidence. This ensures that the generative model becomes a good model of the environment and that the environment complies with the model.

Expected free energy is a way to score alternative policies for planning. This is fundamentally *prospective*—it considers possible future observations—and *counterfactual*—the possible future observations are conditioned on the policies one could pursue. Expected free energy measures the plausibility of action policies relative to preferred (future) states and observations. By scoring policies in terms of their negative expected free energy, creatures engaging in Active Inference effectively believe that they pursue the course of action for which this quantity is lowest. In psychological terms, this implies that a creature's belief about policies directly corresponds to its intention—which it fulfills by acting.

From a conceptual perspective, we can associate minimization of variational free energy and expected free energy with two inferential loops, one nested within the other. Variational free energy minimization is the key (outside) loop of Active Inference, which is sufficient to optimize perception and beliefs about policies. An Active Inference agent can also be endowed with a generative model of the consequences of its action that entails an evaluation of expected free energy (the inside loop). This ability to plan into the future supports prospective forms of action selection by furnishing probability values for policies (Friston, Samothrakis, and Montague 2012; Pezzulo 2012).

2.10 Summary

Active Inference is a theory of how living artifacts underwrite their existence by minimizing surprise—or a tractable proxy to surprise, *variational free energy*—via perception and action. In this chapter, we have sought to motivate this idea starting from a Bayesian treatment of perception as inference and extending this to the domain of action. Bayesian inference rests on a generative model of how sensory observations are generated, which encodes (probabilistically) the organism's implicit knowledge of the world—formalized as prior beliefs and the expected outcomes under alternative states and policies.

The specific take of Active Inference forces us to revisit the usual semantics of a prior in Bayesian inference. Expected states are preferred and include the organism's conditions for survival (e.g., niche-specific goal states), whereas their opposite—surprising states—are dis-preferred. In this way, by fulfilling their expectations, Active Inference agents ensure their own survival. Given the important links between the notion of priors and the conditions that undergird an organism's existence, we can also say that in Active Inference, the identity of an agent is isomorphic with its priors. This terminology will become more familiar later in the book.

Note that in this view, *surprise* (or sometimes *surprisal*) is a formal construct of information theory and not necessarily equivalent to a (folk) psychological construct. Roughly, the more the organism's state differs from the prior (which encodes the preferred states), the more it is surprising—hence Active Inference amounts to the idea that an organism (or its brain) has to actively minimize its surprise to stay alive. Under certain conditions, surprise minimization can be construed as the reduction of the discrepancy between the model and the world. More generally, the quantity that is actually minimized in Active Inference is *variational free energy*. Variational free energy is an (upper-bound) approximation to surprise and can be minimized efficiently using chemical or neuronal message passing and information that is available to the organism's generative model.

Importantly, both perception and action minimize variational free energy in complementary ways: by refining their (posterior belief) estimate and by performing actions that selectively sample what is expected. Furthermore, Active Inference also minimizes *expected free energy* by following policies associated with minimal ambiguity and risk. Expected free energy then extends Active Inference to prospective and counterfactual forms of inference. This completes our journey along the low road to Active Inference. In chapter 3, we will travel the high road, which reaches the same conclusion on the basis of first principles and self-organization.