The development of autonomous virtual agents

Karl Friston

The Wellcome Trust Centre for Neuroimaging, University College London, Queen Square, London WC1N 3BG

Correspondence: Karl Friston
Wellcome Trust Centre for Neuroimaging,
Institute of Neurology
Queen Square, London, UK WC1N 3BG
k.friston@.ucl.ac.uk

Keywords: ideal observers · generative models · Bayesian inference · active inference · free energy · virtual agents · meta-Bayesian · prediction · action · perception

ABSTRACT

This commentary considers some of the basic issues in the development of autonomous virtual agents from a rather general and theoretical viewpoint. It is predicated on an understanding of agents as ideal Bayesian observers, which follows in the long tradition of Helmholtzian ideas about how the brain works and subsequent developments in machine learning and computational neuroscience. The aim of this commentary is to define some key aspects of the problem and discuss potential solutions in relation to a series of specific research questions. In what follows, we try to cast the problem in terms of optimisation, which is particularly pertinent from the point of view of evolutionary schemes. The focus will be on maximizing the evidence for an agent's model of his world or, more precisely, minimizing a variational free energy bound on negative model log-evidence or surprise. This has proven to be a useful framework in the computational neurosciences; and can be shown to be a fairly universal explanation for action and perception. Within this setting, the notion of a model (entailed by a subject) and a model of that subject (entailed by an agent observing the subject) is central. Framing the problem in terms of models raises key questions about their nature; particularly their dynamical form and implicit state spaces. A model-based perspective provides many clear answers to these questions. However, there are some key choices that may need to be formulated carefully, particularly in relation to difference between simply modelling the behaviour of a subject and modelling that behaviour under the constraint that the subject is modelling his world. We will focus on this distinction in terms of the difference between *Bayesian* and *meta-Bayesian* modelling of *subject* behaviours by *virtual agents*.

INTRODUCTION

This commentary is meant to be a discussion document that tries to highlight some of the key issues posed by research questions about the development or autonomous virtual agents. These virtual agents try to predict subject responses and help them make decisions (for example, when and how to release personal identifying information on social networking sites). The purpose of this commentary is to set out the basic issues and provide answers to key questions when they exist and highlight other questions that have yet to be resolved. We first consider the basic imperatives of modelling subjects using a variational free energy formulation of optimal behaviour and optimal inferences (modelling of) that behaviour. We will then move on to the nature of the models underlying this inference and finally consider some specific questions in light of these considerations.

MODELLING THE BEHAVIOUR OF OTHERS

A key question, when it comes to modelling the behaviour of others (such as by a virtual agent or surrogate for a subject) is the implicit modelling of a model. This follows because subjects or real world agents can be seen as optimising a model of their world. This means to model a subject one is implicitly trying to fit or invert a model of a model. In other words, the problem is one of meta-modelling. To make this distinction clear, consider the following problem: We have available to us a sequence of action-state pairs defined on some state space over some finite time or interval. These are the data available to the virtual agent and can be regarded as the *observed responses* of a subject to *observed states*, where the subject is treated as a state space model. The observed states could correspond to inputs the subject has sampled and his action corresponds to outputs or responses emitted by the subject. The objective is to model the subject in order to predict unobserved actions given the history of observed actions and states. To solve this problem, one could model the subject in terms of internal or *hidden states* that remember the inputs and produce outputs or action. In other words, we could model the subject as a state space model, mapping from observed inputs (what the subject sees) to output (what the subject does). This is essentially a system identification problem in which one has to optimise the model and its parameters describing the hidden states of the subject. This could be done with some form of Bayesian model inversion or filtering and Bayesian model comparison to select the best model. This would constitute a straightforward Bayesian approach to optimising a virtual agent that could, in a Bayesian setting, use the predictive distribution over the next action to provide guidance for the subject's next choice. However, note that this approach does not exploit our prior beliefs about the subject. In other words, there are no constraints on the model of the subject as a state space model that embody our knowledge that the subject is itself behaving optimally in some sense. So what sort of constraints could be considered?

IDEAL BAYSIAN OBSERVER CONSTRAINTS

One obvious answer is that the subject is himself modelling the inputs and selecting the outputs in a Bayes-optimal fashion. This implies that the subject has his own model of the world that is providing sensory inputs, which he is optimising in relation to some objective function. The most natural choice here is the evidence for the subject's model that is motivated easily by considering perception and action as inference [1,2,3,4]. Our own work in this area formalises this in terms of *active inference* and the minimisation of a variational free energy bound on surprise [5]; where surprise is negative model log-evidence [4]. This provides a generic account of action and perception that is consistent not only with many neurobiological facts but can also be shown to be imperative for any self-organising system immersed in an inconstant world.

This means that we can regard the subject as optimising his internal (generative) model of the world to minimize variational free energy (a function of sensory inputs and a probabilistic representation under his model). At the same time, he is acting to minimise variational free energy, which simply means he is sampling inputs that confirms his hypotheses. This is active inference. In this setting, optimal behaviour is that which conforms to prior beliefs about what will happen next. In the larger context of modelling the subject, this ideal Bayesian assumption means that the form of the state space model of the subject should itself comply with Bayes-optimality principles. This leads to the notion of Bayes-optimal inferences about a system that is itself Bayes-optimal. In other words, it calls for a meta-Bayesian approach that has been described recently in terms of observing the observer [6].

To my mind, use of a Bayesian or meta-Bayesian approach is probably the least resolved of all the questions that have to be answered to progress the development of autonomous virtual agents: On the one hand, the meta-Bayesian constraints on the model of the subject may greatly reduce the model space and finesse problems in searching over all potential models at the meta-modelling level (i.e. optimising the virtual agents model of the subject's model of the world). On the other hand, the level of computational complexity may increase markedly, in that the agent has to perform a Bayesian inversion (as if it were the subject) to produce predicted responses at each iteration of meta-model inversion. Furthermore, to use a comprehensive model of the subject that even approximates a real human being is, obviously, an enormous undertaking. In my view, this issue deserves some thought. Figure 1 tries to illustrate the difference schematically using variational free energy minimisation as an example of approximate Bayes-optimal inference. The variables and equations in this figure are described in more detail in the appendix, which also provides a slightly more technical overview of active inference for discrete state space models.

THE NATURE OF THE MODEL

Irrespective of whether a Bayesian or meta-Bayesian approach is taken, there are some fundamental restrictions on the nature of the model of the subject that can be articulated. First, we are dealing with a dynamical system, where causes have consequences and the subject's observations lead to subsequent responses. This means that one has to use a dynamical (state space) model of the subject (with or without ideal Bayesian constraints) to predict a series of responses. This can either be in continuous time, where the subject is modelled in terms of differential equations or in discreet time using a hidden Markov model over hidden states. The former has proven useful for simulating low level perceptual and behavioural dynamics in computational neuroscience [7]. However, one might guess that the nature of the data available to the virtual agent would make a discrete time and state space model more viable. In other words, the subject is modelled in terms of probability transitions among hidden states where those probabilities are determined by the current input, which is itself determined by the previous behaviour or action. Action itself is a function of the probability distribution over hidden states and can either be predicted directly or prescribed by minimizing variational free energy under the meta-Bayesian approach (see Figure 1).

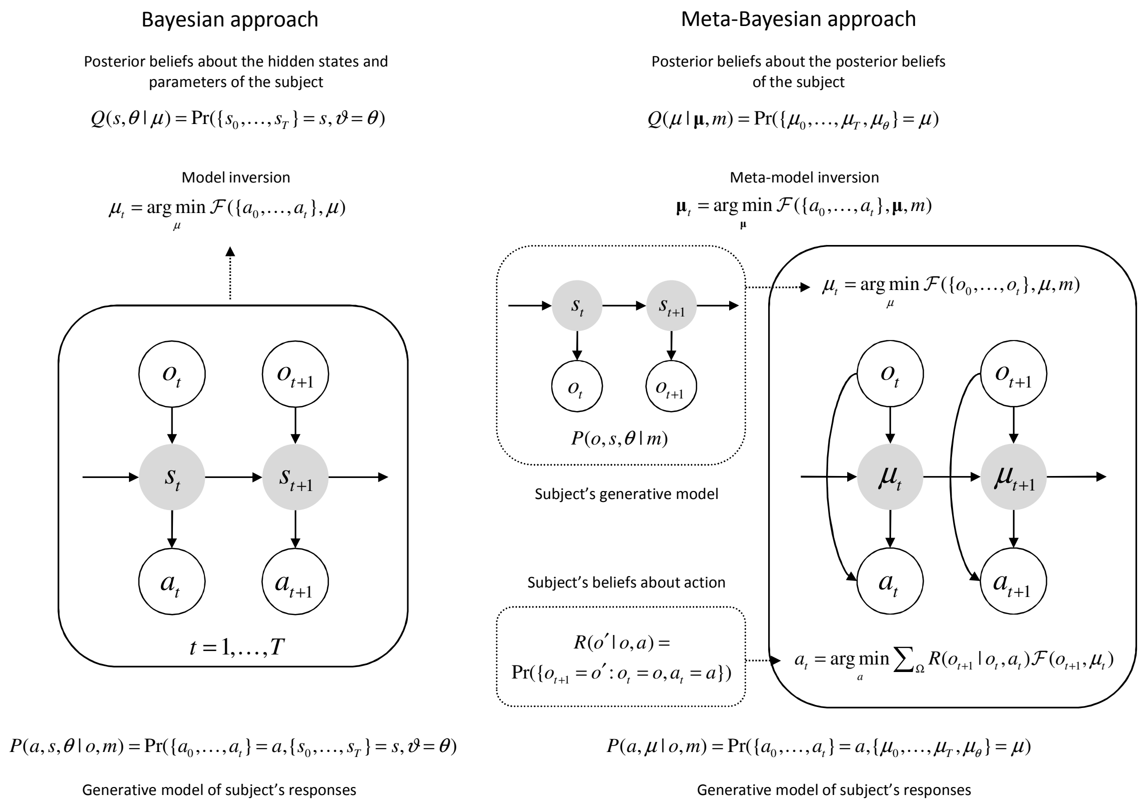

Figure 1: This figure highlights the difference between a Bayesian approach (on the left) and a meta-Bayesian approach (on the right). In this figure, known variables are in white circles and hidden variables are in grey circles. The straightforward Bayesian approach treats the subject as a (Markovian) input-state-output system, where observed states influence hidden states to produce observable responses or action. In this setup, we want to infer the hidden states and the parameters of the transition probabilities among hidden states. Equipped with these, we can then predict subsequent action given any history of observed states. This inference rests on inversion of the implicit generative model of subject responses, which (in this example) uses variational Bayesian procedures to minimise the free energy of observed responses and the sufficient statistics of an approximate posterior distribution over the subject's hidden states. The meta-Bayesian approach is a little more sophisticated and assumes that the hidden states of the subject themselves encode posterior beliefs about what the subject observes. These posterior beliefs are based upon the subject's generative model (shown as an insert) that generates predictions about observed states, given hidden states of the world. These hidden states minimise a variational free energy bound on the evidence for this model of the world, with respect to posterior beliefs. At the same time, actions are selected that minimise the expected free energy of the next of observed state. Crucially, this entails the notion of embodiment because the next observation depends upon the action selected. The hidden states are now posterior beliefs of the subject and meta-model inversion corresponds to evaluating the approximate posterior distribution over the posterior beliefs of the subject. This part is formally the same as in the Bayesian approach but the underlying generative model acknowledges explicitly that the subject is himself performing model inversion (under ideal Bayesian assumptions). Meta-model inversion is numerically more costly; because the hidden states (posterior beliefs of the subject) generating predictions of observed responses are themselves optimising a generative model at each time point. The potential advantage is that the form of the generative model of the subject is specified by the subject's generative model of his world. Please see the appendix for a detailed explanation of the equations and variables.

In either case, it will be necessary to define the input space of observed states, the hidden state space and the action space of observed responses. In addition, model parameters described how inputs affect probability transitions among hidden states and how probabilistic representations of hidden states determine action. One key issue that emerges from these considerations is that the prediction of the subjects final choice (e.g., to release or not personal information or who to share that information with) is only part of the problem. To optimise the model of the subject (in terms of the parameters of the differential equations or probability transition matrices), one clearly needs to use all the observed inputs and subject responses over all preceding time points. In other words, to optimize predictions about a particular behaviour one has to optimise predictions about all responses. This may be a useful consideration when thinking about how to measure and record subject responses to create a space of actions. A related issue here, that is specific to the meta-Bayesian approach, is whether the agent actually thinks about its behaviour or not. This introduces an additional level of complexity; in that to plan or think about behaviour the agents hidden state space (representing hidden states in his world) has to be extended over time [8]. This can create another computational burden and may and may not be necessary for veridical subject modelling.

SOME SPECIFIC QUESTIONS

Given the above observations I will now consider some specific questions:

1) *Would a bottom-up approach allow individual preferences to be represented by an artificial agent?*

A bottom-up approach here is taken to mean the optimisation of a model of the subject and ensuing predictions of their responses. In many respects the above arguments suggest that this is the only approach. Classical top-down approaches can be regarded as heuristics in which one imposing strong prior beliefs about the form and parameters of the subject's model. These (virtual agent) prior beliefs may or may not be right but should certainly be tested against the evidence for these beliefs in empirical behaviour: this is the bottom-up approach.

2) *If so what are the constraints limiting the size of the input domain?*

This will be determined by the computational complexity of model inversion or fitting. These complexity issues could be compounded by the meta-Bayesian approach if the hidden state space of the subject becomes too large (or incorporates future states [8]). One interesting answer this question is that the model is itself determined by the number of hidden states. This means that one can assess the evidence for different models of the subject and optimize the evidence empirically by using Bayesian model selection. Bayesian model selection simply involves quantifying the goodness of a model in terms of its evidence given some data and then selecting the model with the greatest evidence [9]. Practically, the log evidence is usually assessed using variational techniques [10] that avoid stochastic procedures like Gibbs sampling, which can be computationally too burdensome.

Indeed, this is the basis of variational free energy minimisation discussed as a model of a Bayes-optimal subject above.

3) *In order to represent a human how closely does the agent's conceptual structure have to align to humans?*

This question speaks directly to the distinction between the Bayesian and meta-Bayesian approaches above. If the virtual agent's model of the subject can explain the subject's behaviour, then its model must be an approximation or formally equivalent to the model being used by the subject. In this sense, it is absolutely critical to align the agent's conceptual structure (model of the subject) to the actual subject. The degree to which this is possible will clearly depend upon computational resources. Again, even with the Bayesian (as opposed to meta-Bayesian) approach, optimising the model evidence for agent's model will ensure that its conceptual structure and that of the subject will, in some sense, become sufficiently aligned. In this sense, I do not think this is a deep problem.

4) *What function does embodiment having in evolving contextual processing?*

Embodiment enters subtly here as an integral part of the model of the subject. This is because the subject is acting upon the world to disclose his next sensory input. However, the virtual agent is not embodied and does not act upon its world. In this sense, the problem is actually a simple inference problem about an embodied agent. It is not inherently an embodiment problem itself.

5) *What are the specific roles of evolution versus learning?*

From the point of view of Bayesian model selection, there is no difference. The quantities that have to be estimated are the form and parameters of the model of the subject used by the virtual agent to predict subject behaviour. In other words, the virtual agent has to learn the parameters and evolve the model. Both these optimisation processes minimize variational free energy (or maximize model log-evidence). The only difference is that the optimization of the model parameters of a particular model is called *learning*; while *evolution* optimizes the model *per se* (for example, the number and hierarchal deployment of hidden states and allowable state transitions).

6) *Does during the evolutionary phase how closely does an agent's world have to resemble our own or can that be accommodated by an appropriate fitness function?*

From the active inference and meta-Bayesian perspective, the agent's world is basically a world that comprises the subject that is exposed to inputs and produces outputs or responses. As noted above, the model of the subject has to correspond closely to the subject's model or our own model of the world. Having said this, the way that these models are optimised is through the variational free energy associated with each model. This is the appropriate fitness function. In short, the answer to this question is that the appropriate fitness function (variational free energy) ensures the virtual agent's model of the subject world will resemble the subject's model of his world (i.e. our model)

7) *Are current advances in evolutionary algorithms up to the job of reaching these objectives?*

Although this is not my field of expertise, I can see no reason why evolutionary algorithms would not be perfectly suited to optimizing the form or structure of the subject model; where, as noted above, the objective function is the model evidence (or free energy – also known as free fitness [11]).

8) *Can the appropriate dynamical systems that the agent uses be evolved from the bottom up or do we have to create the neural net with the appropriate structure first?*

It is clear that the appropriate dynamical systems (state space models) have to be created from the bottom up. In the context of the current arguments, the process of evolving and optimising this model is the process of creating a neural net with appropriate structure. One can certainly use intuitions and prior beliefs to limit the space of these forms but, conceptually, creating the net is the problem because this is optimising the subject model.

9) *What are the potential implementation deployment obstacles?*

These pertain largely to the sufficient discretisation of state space (under a discrete Markovian state space model of the subject) and sufficient computational resources to invert these models. The key problem here will be an exploration of model space and devising greedy searches that explore appropriate hierarchical forms. Scoring large model spaces is clearly a deep and important problem; however, the evolutionary approach seems ideally suited to this. There are recent advances in the scoring (evaluating the evidence or free energy) of large model spaces that may be useful here [9].

10) *What can be done to overcome these obstacles?*

There are a number of approximate Bayesian inversion techniques that are used in data analysis (and the modelling of perception and action) the most powerful is the use of variational Bayesian procedures (hence the minimization of variational free energy [10]). Although this is beyond my field of expertise there may be some useful pointers in [8], where extremely high dimensional counter-factual state spaces are searched under simplifying (Laplace) assumptions. As noted above, the problem of scoring large numbers of models can also be finessed using the Savage-Dickey ratio and its generalisations [9].

CONCLUSION

In summary, I think the biggest challenge at this stage would be to decide whether to place formal constraints on the state space models of a subject that cast the subject as an ideal Bayesian observer, with or without the capacity to plan (represent future hidden states). Otherwise the notion of using evolutionary schemes to optimise the model of a subject by a virtual agent seems compelling and the most natural approach. I say this because the problem of searching large model spaces and optimising the implicit objective function (free energy or model evidence) is a difficult problem that can probably only be solved using an evolutionary approach. For me, the nature of these models is clear in their broadest terms; however, much will be dictated by the sort of data available to the virtual agent and the ontology of inputs and outputs received by and produced by the subject respectively.

At this stage, I would consider a simple Bayesian system identification approach, where the virtual agent is trying to model the subject as an input-state-output system. This reduces the problem to a conventional system identification problem. If successful, the form of the model (that will include the hidden state space) could be interpreted, *post hoc*, in terms of implicit posterior beliefs and prior expectations held by the subject. The only disadvantage of this approach is that the model is not constrained to include some important features; for example, the subject knows the subspace of sensory inputs that will follow from a particular action and will therefore be able to make predictions that the agent's model of the subject could miss. Note that the agent's model of the subject is not embodied in the sense that the inputs and outputs are all known data-points and the problem is simply the optimisation of a mapping from inputs to outputs. However, from the subject's point of view there is embodiment in the sense that inputs depend upon outputs. The key question here is should the agent's model of the subject use this embodiment as a constraint? The quintessential difference between the Bayesian and meta-Bayesian approaches to system identification reduces to an interpretation of the subject's hidden states as a probabilistic representation (i.e. sufficient statistics or posterior beliefs) of a subject. I hope that these thoughts are useful in framing the discussion of these issues.

Appendix – Active inference and variational free energy

This appendix describes the formalism of active inference, in which the optimisation of action and beliefs about hidden states are treated as two separate processes that maximise model evidence or the marginal likelihood of observations.

The free-energy principle and active inference

The free-energy principle tries to explain how agents occupy a small number of attracting states in terms of minimising the entropy of the (invariant) probability distribution over observed states. This minimisation is assured if agents minimise *surprise* at each time point. Surprise, or more precisely surprisal or self information, is a (probability) measure $-\ln P(o \mid m)$ on the states observed by an agent. Here, surprise is just the negative log likelihood of observations marginalised over hidden states. This marginal likelihood is also known as model evidence. This means that surprise is minimised (approximately or exactly) if agents minimise a variational free energy bound on negative log evidence [6], [7], given a generative model m of state transitions [2], [4].

The free energy principle [4] is based on ergodic arguments about the nature of self-organising systems. These arguments suggest that the long term average of variational free energy upper bounds the (Shannon) entropy of observations over time; which implies that action must minimise variational free energy to resist the second law of thermodynamics (or the dispersion of its states by random fluctuations). This is active inference [5], which extends the minimisation of variational free energy implicit in approximate Bayesian inference on hidden states to include action *per se*. There is a fairly developed literature on free energy minimisation and active inference in the neurosciences; covering things from perceptual categorisation of bird songs, through to action observation [7].

In the present context, active inference unpacks some of the implicit assumptions in Markov decision problems. In particular, it specifies explicitly what the agent knows about the effects of its actions. As we will see, this means that probabilistic transitions among observations are conditioned upon action, which serves to realise posterior beliefs about state transitions. This conditioning of observations on action (as opposed to conditioning states on action) is not unrelated to treatments based on observer operator models and predictive representations of state: see [8] and [9].

Definition: The free energy formulation comprises the tuple $(\Omega, S, A, \vartheta, P, Q, R)$ comprising:

- A finite set of observations Ω.
- A finite set of hidden states S.
- A finite set of actions A.
- Real valued parameters $\vartheta \in \mathbb{R}^d$.
- A *sampling probability* $R(o' \mid o, a) = \Pr(\{o_{t+1} = o' : o_t = o, a_t = a\})$ that observation $o' \in \Omega$ at time $t+1$ follows action $a \in A$, given observation $o' \in \Omega$ at time *t*.
- A *generative probability* $P(o, s, \theta \mid m) = \Pr(\{o_0, \ldots, o_t\} = o, \{s_0, \ldots, s_T\} = s, \vartheta = \theta)$ over observations to time *t*, states at all times and parameters
- A *recognition probability* $Q(s, \theta \mid \mu) = \Pr(\{s_0, \ldots, s_T\} = s, \vartheta = \theta)$ over states at all times and parameters with sufficient statistics $\mu \in \mathbb{R}^d$.

Here, m denotes the form of the generative model or probability $P_m(o, s, \theta) := P(o, s, \theta \mid m)$ and the sufficient statistics of $Q_\mu(s, \theta \mid \mu) := Q(s, \theta \mid \mu)$ encode a probability distribution over a sequence of hidden states $s = \{s_0, \ldots, s_T\}$ and the parameters of the generative probability $\vartheta = \theta$. Crucially, the recognition probability and its sufficient statistics encode hidden states in the future and past, which themselves can change with time: for example, $\mu_k = \{\mu_0^k, \ldots, \mu_T^k\}$, where μ_t^k is the probability over hidden states at time t (in the future or past) under the recognition probability at time k.

Remarks: There are three important distinctions between this setup and that used by classical Markov decision processes (MDPs). As in partially observed MDPs, there is a distinction between states and observations. However, the transition probability over states is replaced by a sampling probability over observations. This means, we can formulate everything in terms of observed states (observations) and inference on hidden states. In other words, the agent does not need to know the effect of its actions on the state of the world. It is instead equipped with a probabilistic mapping between its actions and sensory consequences. This may sound a bit unusual but is a central tenet of active inference, which separates knowledge about the sensory consequences of action from beliefs about the causes of those consequences. In other words, the agent knows that if it moves it will sense movement (cf. proprioception); however, beliefs about hidden states in the world causing movement have to be inferred. These hidden states may or may not include its own action: the key distinction between the *agency free* and *agency based* schemes considered in [8] depends on whether the agent represents its own action or not.

The second distinction is that we have introduced generative and recognition probabilities that are used to infer hidden states. Crucially, these hidden states include future and past states. In other words, the agent represents a sequence or trajectory over states, as opposed to just the current state. Both the generative and recognition probabilities are time-dependent: The generative probability is over the sequence of sensory states up until the current time, while the recognition probability changes with its time-dependent sufficient statistics. This means that the recognition distribution at any one time is over the sequence or trajectory of states at all times. This enables inference about a particular state in the future to change with time. This becomes important when considering planning and agency.

Finally, there are no reward or cost functions. This is an important point and illustrates the fact that active inference does not call upon the notion of reward to optimise behaviour; optimal behaviour minimises variational free energy. In brief, cost functions are replaced by priors over hidden states and transitions, such that costly states are surprising and are avoided by action.

Perception and action

We now require the sufficient statistics of the recognition probability and action to minimise variational free energy

$$\begin{aligned} \mu_t &= \arg\min_{\mu} \mathcal{F}(\{o_0,\ldots,o_t\},\mu) \\ a_t &= \arg\min_{a} \sum\nolimits_{\Omega} R(o_{t+1} \mid o_t, a_t)\mathcal{F}(o_{t+1},\mu_t) \end{aligned} \tag{1}$$

This dual optimisation is usually portrayed in terms of perception and action, by associating the sufficient statistics with internal states of the agent (such as neuronal activity or connection strengths) and associating action with the state of effectors or the motor plant.

By factorising the generative probability $P\left(o,s,\theta \mid m\right) = P(o \mid s,\theta)P(s,\theta \mid m)$ into a likelihood and prior, one can express the free energy in the following three ways:

$$\begin{aligned} \mathcal{F}(o,\mu) &= \mathbf{E}_Q(-\ln P(o,s,\theta \mid m)) - \mathbf{E}_Q(-\ln Q(s,\theta \mid \mu)) \\ &= \mathbf{D}_{KL}(Q(s,\theta \mid \mu) \,||\, P(s,\theta \mid o)) - \ln P(o \mid m) \\ &= \mathbf{D}_{KL}(Q(s,\theta \mid \mu) \,||\, P(s,\theta \mid m)) - \mathbf{E}_Q(\ln P(o \mid s,\theta)) \end{aligned} \tag{2}$$

The first equality in equation (2) expresses free energy as Gibbs energy (expected under the recognition distribution) minus the entropy of the recognition distribution. The second shows that free energy is an upper bound on surprise, because the first (Kullback-Leibler divergence) term is nonnegative by Gibbs inequality. This also means that when free energy is minimised, the recognition density approximates the posterior density $Q(s,\theta \mid o) \approx P(s,\theta \mid o)$ over hidden states and parameters. This formalises the notion of unconscious inference in perception [1], [2] and, under some simplifying assumptions, corresponds to predictive coding [3]. Finally, the last equality expresses free energy as *complexity* (the divergence between the approximate posterior and prior) minus *accuracy* (the expected log likelihood).

The minimisation of free energy, with respect to action in equation (1) is called active inference. This formulation highlights the fact that action selects observable states (not hidden states) that are the least surprising by virtue of having the smallest free energy. The free energy is determined by the sufficient statistics of the recognition distribution. The optimisation of these sufficient statistics (first equality in equation 1) rests upon the generative model and therefore depends on prior beliefs. It is these that specify what is surprising and underwrite Bayes-optimal policies and behaviour.

References

1 Helmholtz, H. Concerning the perceptions in general. In *Treatise on physiological optics*. Dover, New York, 1866/1962.

2 Dayan, P, Hinton, G E, and Neal, RM. The Helmholtz machine. *Neural Computation*, 7 (1995), 889-904.

3 Rao, R P and Ballard, D H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci.*, 2, 1 (1999), 79-87.

4 Friston, K, Kilner, J, and Harrison, L. A free energy principle for the brain. *J Physiol Paris.*, 100, 1-3 (2006), 70-87.

5 Friston, K J, Daunizeau, J, Kilner, J, and Kiebel, S J. Action and behavior: a free-energy formulation. *Biol Cybern.*, 102, 3 (2010), 227-60.

6 Daunizeau, J, den Ouden, H E, Pessiglione, M, Kiebel, S J, Stephan, K E, and Friston, K J. Observing the observer (I): meta-bayesian models of learning and decision-making. *PLoS One*, 5, 12 (Dec 2010), e15554.

7 Friston, K. The free-energy principle: a unified brain theory? *Nat Rev Neurosci.*, 11, 2 (Feb 2010), 127-38.

8 Friston, K, Samothrakis, S, and Montague, R. Active inference and agency: optimal control without cost functions. *Biol Cybernetics* (2012), under review.

9 Friston, K and Penny, W. Post hoc Bayesian model selection. *Neuroimage* , 56, 4 (2011), 2089-99.

10 Beal, M J. Variational Algorithms for Approximate Bayesian Inference. *PhD. Thesis, University College London* (2003).

11 Sella, G and Hirsh, A E. The application of statistical physics to evolutionary biology. *Proc Natl Acad Sci.*, 102 (2005), 9541-6.

12 Feynman, R P. *Statistical mechanics*. Benjamin, Reading MA, 1972.

13 Hinton, G E and van Camp, D. Keeping neural networks simple by minimizing the description length of weights. *Proceedings of COLT-93* (1993), 5-13.

14 Jaeger, H. Observable operator models for discrete stochastic time series. *Neural Computation*, 12 (2000), 1371-98.

15 Littman, M L, Sutton, R S, and Singh, S. Predictive Representations of State. In *Advances in Neural Information Processing Systems* (2002), 1555–61.