What Is Optimal about Motor Control?

Karl Friston[1],*
[1]The Wellcome Trust Centre for Neuroimaging, Institute of Neurology, 12 Queen Square, London WC1N 3BG, UK
*Correspondence: k.friston@ucl.ac.uk
DOI 10.1016/j.neuron.2011.10.018

This article poses a controversial question: is optimal control theory useful for understanding motor behavior or is it a misdirection? This question is becoming acute as people start to conflate internal models in motor control and perception (Poeppel et al., 2008; Hickok et al., 2011). However, the forward models in motor control are not the generative models used in perceptual inference. This Perspective tries to highlight the differences between internal models in motor control and perception and asks whether optimal control is the right way to think about things. The issues considered here may have broader implications for optimal decision theory and Bayesian approaches to learning and behavior in general.

Introduction

Optimal control theory is currently the dominant paradigm for understanding motor behavior in formal or computational terms. It provides a normative model of control that allows many problems to be addressed in a coherent and principled framework (Körding, 2007). Furthermore, it motivates the use of elegant mathematics to solve some difficult problems that the brain contends with (Todorov and Jordan, 2002). The basic premise of optimal control is that optimal movements bring about valuable states. This means that movement can be specified with a value function of states, provided it increases value. Despite the compelling simplicity of this approach, I think it may be wrong for two reasons. First, we know from the physics of flow that motion cannot be specified by a single value function. Second, optimal control theory assumes that movement is caused (determined) by value. However, value is an attribute of states that are caused by movement: it is a consequence, not a cause. This means that the real problem is to understand the acquisition and realization of beliefs that cause movement—in other words, to understand motor control in terms of inference and beliefs. My reading of the recent literature is that there is a shift from the engineering paradigm of optimal control toward a problem formulation in terms of Bayesian inference. However, this paradigm shift may not be complete until we dispense with value functions as the causal explanation of movement. This article compares optimal control and inference and tries to show that inference (1) complies with imperatives that apply to all biological systems, (2) dissolves some hard problems in optimal control, (3) provides a complete specification of control, (4) is neurobiologically plausible, and (5) accounts for action without reference to value. While this may not be important from the point of view of engineering, it may be important for the critical evaluation of optimal control in neuroscience.

Recent developments in motor control theory (Tani, 2003; Verschure et al., 2003; Tani et al., 2004; Jirsa and Kelso, 2005; Wörgötter and Porr, 2005) emphasize sensorimotor dynamics and perceptual inference over conventional optimal control based on forward-inverse models (Miall et al., 1993; Wolpert et al., 1995; Wolpert and Miall, 1996; Todorov and Jordan, 2002; Todorov, 2004; Bays and Wolpert, 2007; Liu and Todorov, 2007; Shadmehr and Krakauer, 2008; Diedrichsen et al., 2010).

See Schaal et al. (2007) for an attempt to reconcile these perspectives. The basic difference is that optimal control assumes that behavior can be reduced to optimizing a value function of states that defines what is optimal. This Perspective focuses on active inference (Friston et al., 2009) as a formal example of the inference approach and compares it with optimal control to ask which of these normative approaches is the most useful. It concludes that optimality may be better understood in terms of prior beliefs about behavior as opposed to value functions. It further shows that active inference resolves several key issues in motor control and unifies current thinking about Bayes-optimal behavior, perception, and learning. Interestingly, similar conclusions follow from arguments based on the equilibrium point hypothesis (Feldman, 2009); namely, there is no need for separate inverse and forward models in motor control because the inverse model can be replaced by (Bayesian) inversion of the forward model. This has no implications for Bayesian formulations of sensorimotor processing (or learning) but has profound implications for notions of optimality, cost functions, and efference copy. We begin with a review of active inference and then consider optimal control schemes.

Active Inference

Active inference is a corollary of the free-energy principle (Friston, 2010) and says that both action and perception minimize surprise. In brief, the motivation for this minimization is to explain how biological systems maintain their biophysical states within bounds and thereby resist the second law of thermodynamics—in other words, to explain how they maintain a homeostasis. They can do this by minimizing the long-term average of surprise, which implicitly minimizes the entropy of their sensory states. Surprise is just the negative log probability of the sensory signals encountered by an agent. In information theory, surprise is called self information, while in statistics it is the negative log model evidence or marginal likelihood. Although agents cannot minimize surprise directly, they can minimize a free energy that is always greater than surprise; hence the free-energy principle. Under some simplifying assumptions, this free energy can be thought of as prediction error. This means that perception can reduce prediction errors by changing predictions (Dayan et al., 1995; Rao and Ballard, 1999), while action reduces prediction errors by changing sensations (Friston et al., 2010). Crucially,
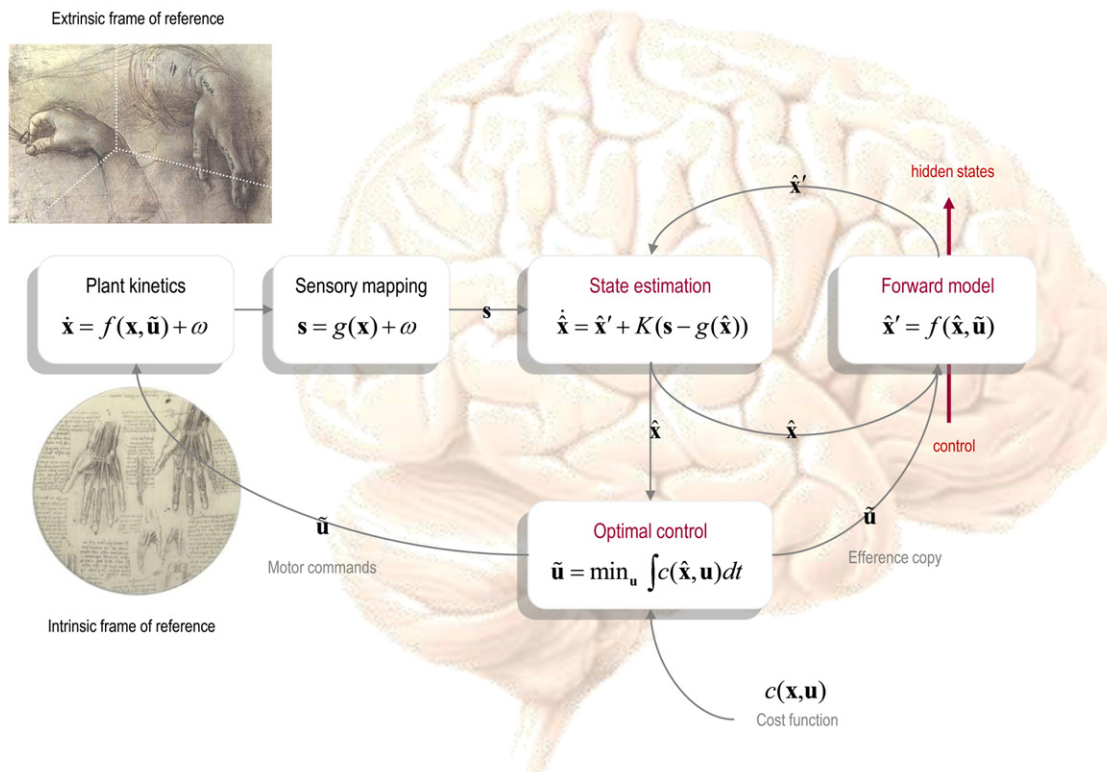
Figure 1. Forward Models in Motor Control

This is a schematic summary of the components commonly found in conventional treatments of optimal motor control. The left side of the figure shows the real world in terms of motor plant kinetics summarized with a stochastic differential equation. The (hidden) states ($\mathbf{x}$) in this equation comprise variables in extrinsic (movement-based) and intrinsic (muscle-based) frames of reference; for example, the motion of the fingertip in Euclidean space and changes in the length of muscle fibers. These states produce exteroceptive (e.g., visual) and proprioceptive (e.g., stretch receptor) sensations through a sensory mapping. Both the kinetics and sensations are subject to random fluctuations (ω). Sensory input ($\mathbf{s}$) is used for hidden-state estimation, summarized here with an extended Kalman-Bucy filter. This Bayesian filter operates in continuous time and updates state estimates ($\hat{\mathbf{x}}$) using predicted motion from a forward model and prediction error ($\mathbf{s} - g(\hat{\mathbf{x}})$) weighted by something called the Kalman gain (K). The prediction error is the difference between sensory input and predictions of that input, $g(\hat{\mathbf{x}})$, given the state estimates. The state estimates are used for optimal control, which returns some control variables (e.g., motor commands) that minimize future cost or loss, specified by a cost function $c(\mathbf{x},\mathbf{u})$, under optimal control ($\tilde{\mathbf{u}}$). This is alternatively referred to as an inverse model that maps from desired trajectories (in an extrinsic frame of reference) to controlled changes in the state of muscles (in an intrinsic frame of reference). Optimal control signals are then sent to the motor plant and (through an efference copy) to the forward model. The forward model then computes the predicted change in hidden states. These predicted changes are integrated with sensory prediction errors by the Kalman-Bucy filter. In this scheme, the forward model can be regarded as a mapping from control to changes in hidden states. Effectively, its role is to finesse the problem of inferring states and thereby optimize control signals. This is necessary because delays and noise on sensory signals could easily confound the implicit closed-loop control used by this scheme.

sensations include both exteroceptive and proprioceptive modalities. This leads to a view of perception as predictive coding and action as the discharge of motor neurons to cancel proprioceptive prediction errors through classical reflex arcs. In this framework, top-down (corticospinal) projections are not motor command signals per se but are predictions about proprioceptive or kinesthetic sensations.

In what follows, we will derive active inference from optimal control theory to identify those components of optimal control that are necessary and those that are not. Optimal control can be cast as active inference with three simplifications: the first formulates optimal control in terms of predictive coding, the second replaces optimal control with motor reflex arcs, and the third replaces value functions with prior beliefs. The first simplification provides a unifying perspective on perception and action and highlights the central role of Bayesian filtering in model inversion. Furthermore, it shows that forward models

in motor control are not the generative models that are actually inverted. The second simplification finesses the problem of delays in descending signals and reinstates classical reflex arcs as an integral part of motor control. Finally, the replacement of value and cost functions with prior beliefs about movements removes the optimal control problem completely.

Conventional Motor Control Schemes and Active Inference

Figure 1 is based on a nice overview of conventional schemes by Frens and Donchin (2009). This schematic tries to accommodate the key ingredients of optimal control, ranging from early notions about Smith predictors (Miall et al., 1993) to the more recent synthesis of optimal control and state estimation (Todorov, 2004; Körding and Wolpert, 2004; Paulin, 2005). Figure 1 uses a nonlinear formulation in continuous time to emphasize that these schemes have to be realized neurobiologically. The three
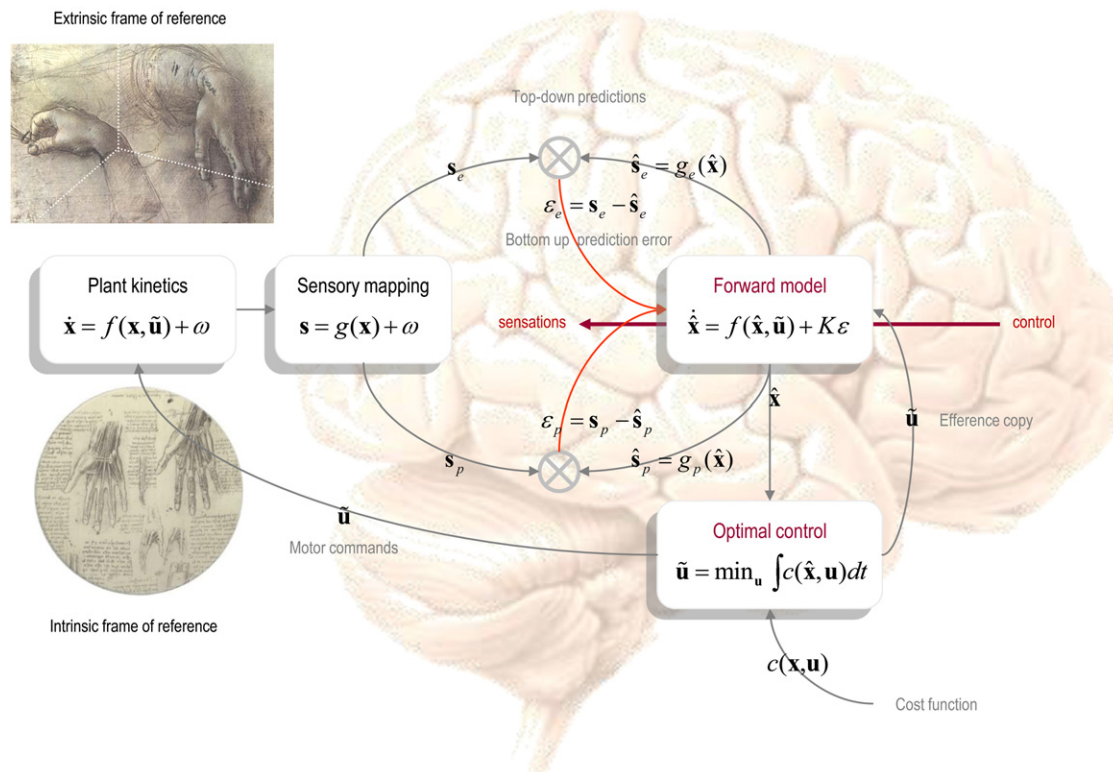
Figure 2. Predictive Coding in Motor Control
This represents the same scheme as in the previous figure, but here state estimation has been absorbed into the forward model, and the prediction errors are represented explicitly. Furthermore, we have made a distinction between exteroceptive (ε_e) and proprioceptive (ε_p) prediction errors reporting on hidden states in extrinsic and intrinsic frames of reference, respectively. These prediction errors are simply the difference between the sensory input observed ($\hat{\mathbf{s}}$) and predicted ($\hat{\mathbf{s}} = g(\hat{\mathbf{x}})$). The resulting scheme now looks like the scheme employed by predictive coding, which also rests on Bayesian (Kalman-Bucy) filtering. In this form, top-down predictions from the forward model are compared with sensory inputs to produce bottom-up prediction errors (red connections) that enter the Bayesian filter. Crucially, the mapping from hidden states to sensations is now part of the forward (generative) model.

key components are (1) an inverse model or optimal control, (2) a forward model, and (3) state estimation (see figure legend). In brief, the optimal control computes command signals that minimize some cost function, specifying the desired movement. Although this seems straightforward, it assumes that an underlying optimality equation can be solved (Bellman, 1952). This is a difficult problem with several approximate solutions, ranging from backward induction to dynamic programming and reinforcement learning (Sutton and Barto, 1981). Optimal control signals depend on the (hidden) states of the motor plant that are estimated using sensory signals. This estimation is generally construed as a form of Bayesian filtering, represented here with a (continuous time) Kalman-Bucy filter. Here, filtering means estimating hidden states from a sequence of sensory observations in a Bayes-optimal fashion. This involves supplementing predicted changes with updates based on sensory prediction errors. The predicted changes are the outputs of the forward model, based on state estimates and optimal control signals. This requires the controller to send an efference copy of its control signals to the forward model. In this setup, the forward model can also be regarded as finessing state estimation by supplementing noisy (and delayed) sensory prediction errors with predictions to provide Bayes-optimal state estimates.

Crucially, these estimates can finesse problems incurred by sensory delays in the exchange of signals between the central and peripheral nervous systems.

In summary, conventional schemes rest on separate inverse and forward models, both of which have to be learned. The learning of the forward model corresponds to sensorimotor learning, which is generally considered to be Bayes optimal. Conversely, learning the inverse model requires some form of dynamic programming or reinforcement learning and assumes that movements can be specified with cost functions that are supplied to the agent.

Predictive Coding and Motor Control
Figure 2 shows a minor rearrangement of the conventional scheme to highlight its formal relationship with predictive coding. Mathematically, the predicted changes in hidden states have been eliminated by substituting the forward model into the state estimation. This highlights a key point: the generative model inverted during state estimation comprises the mapping between control signals and changes in hidden states and the mapping from hidden states to sensory consequences. This means that the forward model is only part of the full generative model implicit in these schemes. Furthermore, in Figure 2, sensory prediction errors are represented explicitly to show how their construction

corresponds to predictive coding. In predictive coding schemes, top-down predictions are compared with bottom-up sensory information to create a prediction error. Prediction errors are then passed forward to optimize predictions of the hidden states, shown here using the Kalman-Bucy filter. There is a large literature on predictive coding as a model of perceptual inference, which is considered to be a biologically plausible form of Bayesian filtering (Mumford, 1992; Rao and Ballard, 1999; Friston et al., 2006). Note that the sensory prediction errors in predictive coding (Tseng et al., 2007; Wei and Körding, 2009) have nothing to do with reward prediction errors in optimal control and reinforcement learning (Schultz and Dickinson, 2000; Gläscher et al., 2010). Sensory prediction errors are required for online state estimation (inference) and optimizing (learning) the forward model. Conversely, reward prediction errors are concerned solely with learning the inverse model, in terms of value functions or cost-to-go (the path integral of cost under optimal control). Reward prediction errors are generally invoked in the context of reward learning; however, exactly the same errors are required when learning the cost-to-go in motor control.

In summary, it is straightforward to cast optimal motor control in terms of predictive coding. In this setting, the forward model is part of a generative model mapping from control to sensory consequences. This distinction may be trivial from the perspective of optimal control schemes, but it is important for active inference, as we will see.

Figure 2 distinguishes between exteroceptive and proprioceptive prediction errors on sensations caused by (hidden) states in extrinsic and intrinsic frames of reference. Here, the (high-dimensional) intrinsic frame contains the state of the motor plant (e.g., muscle fibers). Conversely, the (low-dimensional) extrinsic frame contains movement in extrapersonal space (e.g., a head-centered frame of reference). Intrinsic and extrinsic frames are used in the sense of Kakei et al. (2003) and Shipp (2005): Kakei et al. discuss movement representations in terms of the coordinate transformations that begin with an "extrinsic coordinate frame representing the spatial location of a target and end with an intrinsic coordinate frame describing muscle activation patterns." In Feldman and Levin (1995), these frames of reference are considered in terms of physical (intrinsic) and action-perception (extrinsic) frames. The distinction is important because optimal control has to invert a mapping from (1) control signals to consequences in an intrinsic (muscle-based) frame and then (2) from an intrinsic to an extrinsic (movement-based) frame in which desired movement is defined. In short, the inverse mapping comprises two parts: from an extrinsic to an intrinsic frame and from an intrinsic frame to control signals. The second part of the inversion is easy because there is a simple relationship between motor neuron activity and its consequences (if an alpha motor neuron fires, its extrafusal muscle fibers contract). However, the first part makes inversion difficult because there are many intrinsic degrees of freedom that interact to produce a trajectory in extrinsic coordinates. In what follows, we will separate the easy (intrinsic) and hard (extrinsic) inverse problems and then dispense with the hard problem.

Reflex Arcs and the Easy Inverse Problem
Recall that the motivation for state estimation in optimal control is to finesse problems with noisy and delayed sensory input.

However, there are also delays in descending control signals from the motor cortex. These can be discounted if we consider classical reflex arcs to be solving the easy (intrinsic) inverse problem. In other words, if motor neurons are wired to suppress proprioceptive prediction errors in the dorsal horn of the spinal cord, they effectively implement an inverse model, mapping from desired sensory consequences to causes in intrinsic (muscle-based) coordinates. In this simplification of conventional schemes, descending motor commands become top-down predictions of proprioceptive sensations conveyed by primary and secondary sensory afferents. Note that this is not an open-loop scheme, because top-down predictions are part of a closed loop that optimizes estimates of hidden states using bottom-up (e.g., visual) sensations.

This simplification speaks to the recursive and hierarchical anatomy of the motor system (Grafton and Hamilton, 2007; Shipp, 2005) and acknowledges the role of nested, closed-loop dynamics at both peripheral and central levels. In this scheme, optimal control signals prescribe action indirectly through predictions about desired proprioceptive consequences. This means that their role is to provide predictions about changes in hidden states that minimize cost. These predictions (from the forward model in Figure 1) require optimal control to solve the hard (extrinsic) inverse problem. However, this is no longer necessary because control signals are not required in intrinsic coordinates (because the intrinsic consequences of extrinsic predictions drive action). It is therefore sufficient to provide the forward model with predictions about desired trajectories in an extrinsic frame of reference. This means that we do not have to solve the hard problem of working out how (intrinsic) muscle contractions produce (extrinsic) movements; we only have to solve the forward problem of how (extrinsic) movements stretch (intrinsic) muscles. In other words, the inverse model (optimal control) is unnecessary. This brings us to active inference.

Active Inference, Cost, and Priors
Active inference eschews the hard inverse problem by replacing optimal control signals that specify muscle movements (in an intrinsic frame) with prior beliefs about limb trajectories (in an extrinsic frame). The resulting scheme is shown in Figure 3, where the forward model now maps from prior beliefs about desired trajectories to their sensory consequences. This model is formally identical to hierarchical models used for perceptual inference. Here, motor commands become descending predictions of proprioceptive sensations, while their exteroceptive homologs become corollary discharges (see left panel of Figure 4). In short, with one simple manipulation, we have eliminated the need for optimal control and the intractable solution of the Bellman optimality equation. This changes the normative model of motor control fundamentally: optimal control relies on an inverse model to provide control signals that prescribe trajectories that are optimal in relation to some cost function. In active inference, the trajectories are Bayes optimal (in relation to sensory evidence or free energy), and there is no inverse model or cost function. This is important because Bayes-optimal trajectories do not necessarily have well-defined cost functions (see below). In short, active inference is consistent with Bayesian perception and sensorimotor learning of generative forward
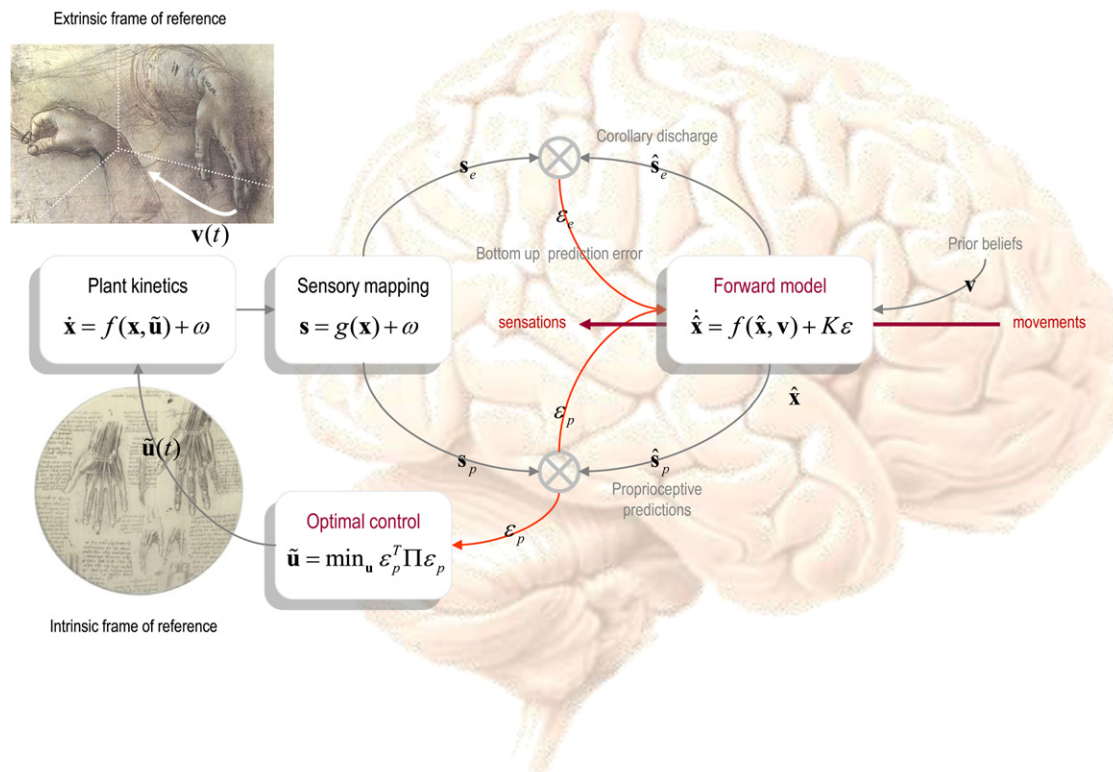
Figure 3. Active Inference
This figure represents the final simplification of the predictive coding scheme of the previous figure. Here, cost functions have been replaced by prior beliefs about (desired) trajectories in an extrinsic frame of reference. These beliefs enter the Bayesian filter to guide predictions of sensory inputs. Proprioceptive predictions are fulfilled in the periphery through classical motor reflex arcs, while predictions of exteroceptive inputs correspond to corollary discharge and are an integral part of perceptual inference. Note that optimal control now reduces to simply suppressing proprioceptive prediction errors. This is active inference.

models and removes the problem of computing the cost-to-go. This is summarized nicely in Feldman (2009): "Efference copy-based and internal model theories consider a problem of a mapping between desired movements and associated motor commands. It is assumed that this problem is solved by pre-programming of the requisite commands with the help of inverse and forward internal models. In contrast, by utilizing frames of reference as action-producing tools, the system does not need to program these commands."

It should be noted that there is no free lunch when replacing cost functions with prior beliefs. It is well known that the computational complexity of a problem is not reduced when formulating it as an inference problem; see Littman et al. (2001) for a treatment of this in the setting of stochastic satisfiability problems. This fact is evidenced by the many procedures that are found in both approximate optimal control and Bayesian inference. Examples here include minimization of Kullback-Leibler divergences (Todorov, 2008; Kappen et al., 2009) and expectation maximization (Toussaint and Storkey, 2006), both of which can be formulated as minimizing free energy (Neal and Hinton, 1998). In one sense, active inference replaces a hard optimal control problem with a hard inference problem. Having said this, the nice thing about active inference is that these problems can be solved in a simple and neurobiologically plausible fashion: by effectively equipping predictive coding schemes with classical reflex arcs (see Figure 4 and Mumford, 1992; Friston, 2008). Perhaps the most definitive argument in favor of active inference, as a normative model of motor control, is that prior beliefs about behavior emerge naturally as top-down or empirical priors during hierarchical perceptual inference. This contrasts with optimal control, which, at the end of the day, still has to explain how cost functions themselves are optimized. In short, active inference eliminates the homunculus implicit in cost functions.

Active Inference and Optimal Control
In this section, we compare and contrast active inference with optimal control at a number of different levels. We will start at the level of implementation and frames of reference and then turn to relationships at the theoretical level, in terms of the duality between cost functions and priors and between optimal control and inference.

In conventional schemes, the intrinsic frame of reference contains the causes (changes in muscle length), while the consequences (changes in limb position) are in extrinsic coordinates. Active inference turns this on its head and regards prior beliefs that cause movement to exist in an extrinsic frame, while the consequences unfold in intrinsic coordinates. In what sense are these perspectives equivalent? Intuitively, one can either regard a limb as being pulled by a muscle or the muscle as being
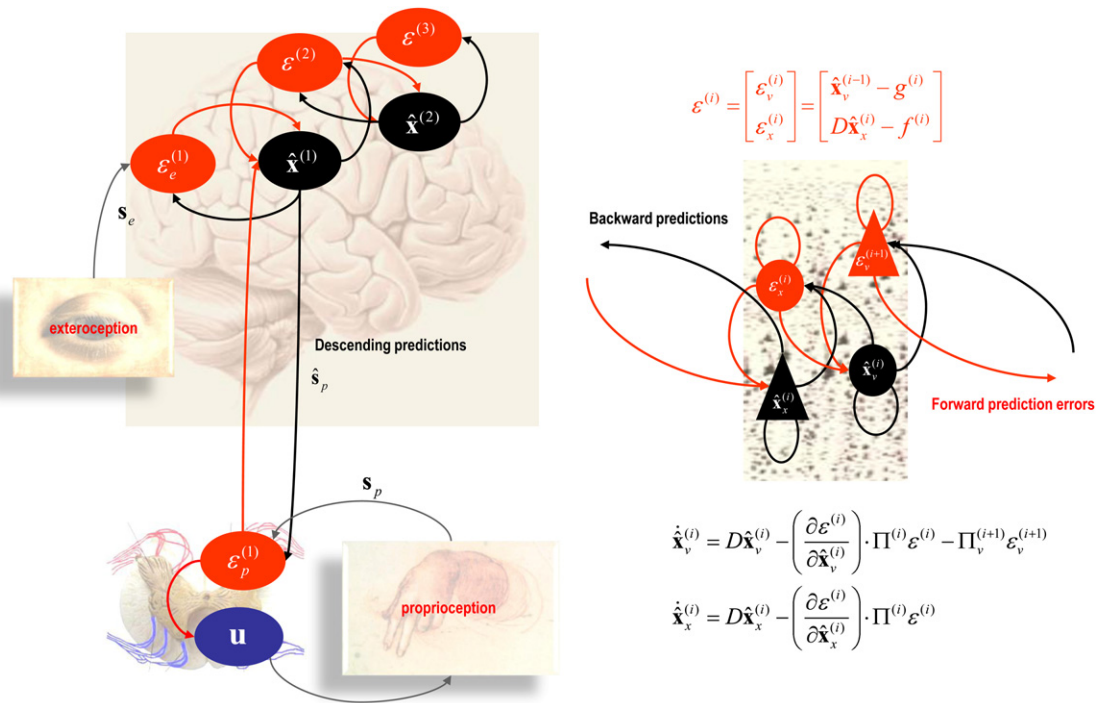
Figure 4. Hierarchical Message Passing in the Brain
This figure illustrates the sort of neuronal architecture that might implement active inference. The left panel shows a schematic of predictive coding schemes in which Bayesian filtering is implemented by neuronal message passing between superficial (red) and deep (black) pyramidal cells encoding prediction errors and conditional predictions or estimates, respectively (Mumford, 1992). In these predictive coding schemes, top-down predictions conveyed by backward connections are compared with state estimates at the lower level to form a prediction error. This prediction error is then passed forward to update the state estimates in a Bayes-optimal fashion. In active inference, this scheme is simply extended to include classical reflex arcs, where proprioceptive prediction errors drive alpha motor neurons in the ventral horn of the spinal cord to elicit extrafusal muscle contractions and changes in primary sensory afferents from muscle spindles. These suppress prediction errors encoded by Renshaw cells. The right panel presents a schematic of prediction error and state estimation units at some arbitrary level in a cortical hierarchy. In this example, there is a distinction between hidden states ($\mathbf{x}_x$) that model dynamics and hidden causes ($\mathbf{x}_v$) that mediate the influence of one level on the level below. The equations correspond to a generalized Bayesian filtering or predictive coding in generalized coordinates of motion, as described in (Friston, 2008). In this hierarchical form, $f^{(i)} := f(\mathbf{x}_x^{(i)}, \mathbf{x}_v^{(i)})$ corresponds to the equations of motion at the ith level, while $g^{(i)} := g(\mathbf{x}_x^{(i)}, \mathbf{x}_v^{(i)})$ links levels. These equations constitute the agent's prior beliefs. D is a derivative operator, and $\Pi^{(i)}$ represents precision or inverse variance. These equations were used for the simulations presented in the next figure.

pushed by the limb. However, from the point of view of hidden states (muscle length and limb position), the two scenarios are identical. In other words, the semantics of push versus pull are purely heuristic; the underlying trajectories (in both frames of reference) are simply solutions to the appropriate Euler-Lagrange equations of motion. In active inference, movements caused by changes in muscle length are modeled as movements that cause changes in muscle length; cf. the Passive Motion Paradigm (Mussa Ivaldi et al., 1988). Intuitively, this makes sense in that we are aware of movements, not muscles.

Can every movement specified by a cost function also be specified by a prior belief? An equivalence between cost functions and prior beliefs can be established by appealing to the complete class theorem (Brown, 1981; Robert, 1992). This states that any behavior is Bayes optimal for at least one prior belief and cost function. However, this pair is not necessarily unique, which means that one can exchange prior beliefs and cost functions to produce the same motor behavior. This is exploited in active inference to provide a biologically plausible solution to the motor control problem that can be regarded as a predictive coding with motor reflexes. This scheme can also be regarded as an instance

of the equilibrium point hypothesis (Feldman and Levin, 1995), in which fixed points are replaced by trajectories that are specified by prior beliefs about motion. In active inference, these are actually empirical priors that are continuously updated during the perceptual inversion of hierarchical generative models. In this setting, the optimal trajectory is just the movement that has the greatest posterior probability, given the current context. See Figure 4.

Optimal Control as Inference
The duality between optimal control and estimation has been clearly articulated by Todorov (2008) and dates back to the inception of Kalman filtering. This equivalence was exploited by early proposals to replace cost with an auxiliary random variable conditioned on a desired observation. This means that minimizing cost is equivalent to maximizing the likelihood of desired observations (Cooper, 1988; Pearl, 1988; Shachter, 1988). Subsequent work focused on efficient methods to solve the ensuing inference problem (Jensen et al., 1994; Zhang, 1998). Later, Dayan and Hinton (1997) proposed an expectation maximization algorithm for reinforcement learning in the case of immediate rewards, while Toussaint and Storkey (2006) cast the

problem of computing optimal policies as a likelihood maximization problem. More recently, variational Bayesian procedures have been applied to optimal decision-making problems in Markov decision processes (Botvinick and An, 2008; Hoffman et al., 2009; Toussaint et al., 2008) and stochastic optimal control (Mitter and Newton, 2003; Kappen, 2005; van den Broek et al., 2008; Rawlik et al., 2010). These approaches appeal to variational techniques to provide efficient and computationally tractable solutions, in particular by formulating the problem in terms of Kullback-Leibler minimization (Kappen, 2005) and path integrals of cost functions using the Feynman-Kac formula (Theodorou et al., 2010; Braun et al., 2011). So what does active inference bring to the table?

Prior Beliefs or Cost Functions?

Active inference goes beyond noting a formal equivalence between optimal control and Bayesian inference. It considers optimal control a special case of inference in the sense that there are policies that can be specified by priors that cannot be specified by cost functions. This follows from the fundamental lemma of variational calculus, which says that that a policy or trajectory has both curl-free and divergence-free components, which do and do not change value, respectively. This means that value can only specify the curl-free part of a policy. A policy or motion that is curl free is said to have detailed balance and can be expressed as the gradient of a Lyapunov or value function (Ao, 2004). The implication is that only prior beliefs can prescribe divergence-free motion of the sort required to walk or write. This sort of motion is also called solenoidal, like stirring a cup of coffee, and cannot be specified with a cost function, because every part of the trajectory is equally valuable. So why is this not a problem for active inference?

The difference between active inference and optimal control lies in the definition of value or its complement, cost-to-go. In optimal control, value is the path integral of a cost function, whereas in active inference, value is simply the log probability or sojourn time a particular state is occupied under prior beliefs about motion. This sort of value does not require cost functions. Technically speaking, in stochastic optimal control, action is prescribed by value, which requires the solution of something called the Kolmogorov backward equation (Theodorou et al., 2010; Braun et al., 2011). This equation is integrated from the future to the present, starting with a cost function over future or terminal states. Conversely, in active inference, action is prescribed directly by prior beliefs, and value is determined by the stationary solution of the Kolmogorov forward equation (Friston, 2010; Friston and Ao, 2011). See Mitter and Newton (2003) for a discussion of forward and inverse Bayes formulae and their variational characterizations in terms of optimality. The forward type of optimality in active inference is closely related to the optimality introduced recently for the control of stochastic nonlinear problems with solenoidal or periodic motion, such as in locomotion, in which "the stationary state-distribution of the optimally-controlled process" is approximated (Tassa et al., 2011). In short, optimal motion is determined by prior beliefs, which endow states with a particular value; however, value is a consequence, not a cause, of optimal behavior. The crucial thing here is that cost-to-go and surprise are the same thing. This ensures that maximizing the long-term average of value is

the same as minimizing the entropy of sensory states. This is mandated by the free-energy principle and is the same as maximizing Bayesian-model evidence. Both value and surprise are optimized by Bayesian inference, but neither depends on cost functions. We will see an example of cost-free optimality below.

In summary, the tenet of optimal control lies in the reduction of optimal motion to flow on a value function, like the downhill flow of water. Conversely, in active inference, flow is specified directly in terms of equations of motion that constitute prior beliefs, like patterns of wind flow. The essential difference is that prior beliefs can include solenoidal flow (e.g., atmospheric circulation, or the Coriolis Effect) that cannot be specified with (scalar) value functions. Having said this, I do not want to overstate the short-comings of optimal control in specifying limit cycle or solenoidal motion; for example, there are compelling examples in the recent literature on simulated walking (Wang et al., 2009). These schemes employ simultaneous trajectory optimization, which uses an explicit representation of the trajectory (as opposed to sequential algorithms that only represent the action sequence) (Kameswaran and Biegler, 2006). This generalization replaces cost functions of a particular state with a cost function over trajectories. Effectively, this converts the problem of optimizing a sequence of movements into optimizing a value function on a high-dimensional state space, whose coordinates are states at different times. A point in this space encodes a sequence or trajectory. However, this begs the question of how one would specify an itinerant sequence of sequences, without invoking even higher-dimensional representations of state space. This is accommodated easily in inference, in which prior beliefs about sequences of sequences are encoded directly by hierarchies of attractors or central pattern generators (Kiebel et al., 2008). Another generalization of optimal control is to consider value functions that change with time (Todorov and Jordan, 2002). Intuitively, this would be like guiding a donkey with a moving carrot (as opposed to placing the carrot at a fixed location and hoping the donkey finds it). However, this just replaces questions about the donkey with questions about how the carrot moves. In active inference, the carrot can be regarded as prior beliefs (that specify the desired trajectory), while the donkey is compelled by posterior beliefs and classical reflexes to follow the carrot.

Efference Copy and Corollary Discharge

Finally, active inference provides a particular interpretation of efference copy (EC) and corollary discharge that predicts the sensory consequences of descending motor signals. In active inference, descending signals are in themselves predictions of sensory consequences (cf. corollary discharge). In this sense, every backward connection in the brain (that conveys top-down predictions) can be regarded as corollary discharge, reporting the predictions of some sensorimotor construct. The fact that high-level (amodal) representations have both motor and sensory consequences highlights the intimate relationship between action and perception. Note that efference copy per se disappears in active inference. This may not be too surprising, given the assertion that the "solutions to the three classical problems of action and perception (the posture-movement problem, problems of kinesthesia, and visual space constancy) offered by the EC theory in particular or by the internal model theory in general are physiologically unfeasible" (Feldman, 2009).
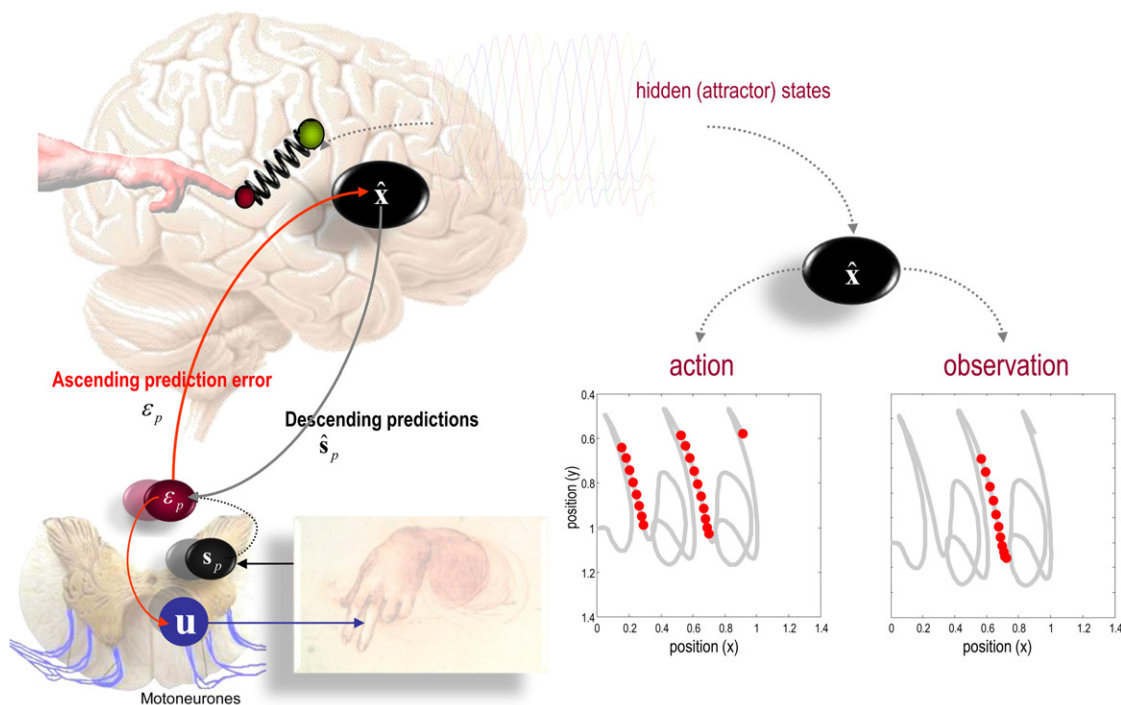
Figure 5. Active Inference and Action Observation

This schematic summarizes the results of the simulations of action observation reported in Friston et al. (2011). The left panel pictures the brain as a forward or generative model of itinerant movement trajectories (based on a Lotka-Volterra attractor, whose states are shown as a function of time in colored lines). This model furnishes predictions about visual and proprioceptive inputs, which prescribe movement through reflex arcs at the level of the spinal cord (inset on the lower left). The variables have the same meaning as in the previous figures. The mapping between attractor dynamics and proprioceptive consequences is modeled with Newtonian mechanics on a two-jointed arm whose extremity (red ball) is drawn to a target location (green ball) by an imaginary spring. The location of the target is prescribed (in an extrinsic frame of reference) by the currently active state in the attractor dynamics. These attractor dynamics and the mapping to an extrinsic (movement) frame of reference constitute the agent's prior beliefs. The ensuing posterior beliefs are entrained by visual and proprioceptive sensations by prediction errors during the process of inference, as summarized in the previous figure. The resulting sequence of movements was configured to resemble handwriting and is shown as a function of location over time on the lower right (as thick gray lines). The red dots on these trajectories signify when a particular neuron or neuronal population encoding one of the hidden attractor states was active during action (left) and observation of the same action (right). More precisely, the dots indicate when responses exceeded half the maximum activity and are shown as a function of limb position. The left panel shows the responses during action and illustrates both a place-cell-like selectivity and directional selectivity for movement in an extrinsic frame of reference. The equivalent results on the right were obtained by presenting the same visual information to the agent but removing proprioceptive sensations. This can be considered to be a simulation of action observation and a mirror of neuron-like activity. This is an interesting example from the point of view of the current discussion, because it highlights the intimate relationship between perceptual inference and action.

Discussion

The arguments above are presented in a rather abstract way, without substantiating the assumptions or background on which active inference rests. This omission is probably best addressed by reference to work showing that cost functions and optimal policies can be formulated as prior beliefs in the context of active inference (Friston et al., 2009) and that the same scheme can be extended to include heuristic policies (Gigerenzer and Gaissmaier, 2011) formulated using dynamical systems theory (Friston, 2010). In the motor domain, active inference provides a plausible account of retinal stabilization, oculomotor reflexes, saccadic eye movements, cued reaching, sensorimotor integration, and the learning of autonomous behavior (Friston et al., 2010). In this context, Bayes-optimal sensorimotor integration (Körding and Wolpert, 2004) is an emergent property that is mandated by absorbing action into perceptual inference. This is illustrated nicely when simulating action observation. An example is provided in Figure 5, in which the same scheme is used to generate autonomous (handwriting) movements and to recognize the same movements when performed by another

agent. The equations used in this example can be found in Friston et al. (2011). This example was chosen to show that the same (neuronal) representations play the role of prior beliefs during the prosecution of an action and recognizing the same action when observed. In this sense, the very existence of mirror neurons (that respond selectively to actions and observation of the same action) are an empirical testament to the duality between optimality and inference. It would be interesting to see whether this simulation of the mirror neuron system could be reproduced using optimal control theory (Miall, 2003). This is a slightly disingenuous challenge because optimal control cannot reproduce handwriting as a result of requisite motion being solenoidal. As noted above, this is a shortcoming of optimal control when it comes to itinerant (sequential and wandering) movements. In short, the compete class theorem suggests that any optimal trajectory specified by a cost function can be specified by a prior belief but that not every optimal trajectory can be specified by a cost function.

The issues addressed in this review are largely theoretical in nature and speak to formal or computational modeling of motor

control: specifically, should these models be based on optimal control theory or optimal Bayesian inference. However, the answer has some profound neurobiological implications. For example, if descending motor commands are top-down predictions, then descending motor efferents should share physiological and anatomical characteristics with top-down or backward connections in other systems. Indeed, descending projections from primary motor cortex share many features with backward connections in visual cortex: they originate in infragranular layers and target cells expressing NMDA receptors. This is somewhat paradoxical, from the orthodox perspective (Shipp, 2005), because backward modulatory characteristics (Sherman and Guillery, 1998) would not be expected of driving motor command signals. This apparent paradox is resolved by active inference, which also provides a principled explanation for why the motor cortex is agranular (R. Adams, personal communication).

There are clearly many operational issues that attend the distinction between optimal control and active inference. For example, how does active inference compensate for altered limb dynamics or external perturbations? A treatment of this can be found in Friston et al. (2010), in which movement trajectories are shown to be remarkably robust to perturbations, both to forces on a limb and fluctuations in motor gain. Heuristically, active inference counters unpredicted forces immediately (to suppress prediction errors on force); in contrast, optimal control can only adjust its (state-dependent) control signals after unpredicted forces change the state of the motor plant. Another key area we have not considered is the learning or acquisition of prior beliefs. In optimal control, the value function is learned, whereas in active inference, the problem reduces to learning the parameters (of the equations of motion) that constitute prior beliefs. This is a standard problem in inference and corresponds to perceptual learning. For example, the agent depicted in Figure 5 could optimize its parameters during action observation (with respect to free energy) and use them to reproduce observed behavior during action. Note that this form of imitation learning reduces to pure perceptual learning and eschews the inverse optimal control problem of inferring the value function by observing an optimally controlled system (Dvijotham and Todorov, 2010).

From the point of view of active inference, cost functions represent one particular way of specifying prior beliefs about the future. It is interesting to speculate that their formal simplicity makes them an attractive candidate for representations of goals at a cognitive level. In other words, in the conscious control of behavior, we may represent cost functions explicitly. This is implicit in the use of optimal decision theory to describe planning and choices (Botvinick and An, 2008; Gläscher et al., 2010). However, the arguments presented here suggest that cost functions per se are not an inherent part of motor control, because they can only specify the component of movement trajectories with detailed balance.

Conclusion

In summary, active inference is appealing for several reasons. First, it dispenses with optimal control (in the sense of solving optimality equations and learning cost-to-go). This is important because there are no biologically plausible schemes that can handle nonlinear (and divergence-free) control problems in continuous time. Second, it finesses problems with delayed control signals in classical formulations. In other words, descending corticospinal signals are predictions that are fulfilled at the peripheral level using fast closed-loop mechanisms (i.e., peripheral reflex arcs). These predictions can anticipate delays if they are part of the generative model. Finally, active inference resolves Bernstein's problem (Bernstein, 1967). Bernstein's problem rests on the many-to-one mapping from the intrinsic frame to the extrinsic frame. This induces indeterminacy in producing a particular trajectory. The resulting, ill-posed nature of the inverse problem means that one has to invoke auxiliary objective functions like minimum jerk to provide unique solutions. In active inference, these problems are resolved by prior beliefs about the trajectory (that may include minimal jerk) that uniquely determine the (intrinsic) consequences of (extrinsic) movements.

A forthcoming review of sensorimotor learning (Wolpert et al., 2011) highlights three key challenges for motor control theory, which can be addressed in light of the above discussion:

- "It is not clear whether the learning models developed will generalize to tasks such as tying shoelaces or learning to skateboard." Optimal control theory will fail here because these behaviors (like handwriting) entail solenoidal motion.
- "To date, relative few principles from the study of biological sensorimotor control have found their way into robotics." This may be because the solution of the optimal equations (when they exist) is intractable (or notoriously slow) in realistic settings. It is notable that compelling reproductions of animate movements in robotics (Tani, 2003) can be cast as active inference, in which the inverse model (optimal control) is replaced by model inversion.
- "Although significant progress has been made in computational sensorimotor control, the field has been less successful in linking computational models to neurobiological models of control." This may be because inverse and forward models do not exist, because there is no optimal controller. An important corollary of this is that optimal control schemes require both the forward model and inverse model to be learned (through use-dependent learning and value learning, respectively). In active inference, there is only use-dependent learning of the generative model.

This essay poses a provocative question about the usefulness of optimal control, inverse models, and cost functions in motor control theory. I half expect an answer of the form, "What you say is interesting, but you have overlooked one fundamental problem," or "Optimal control theory is necessary to explain the following empirical observation." Perhaps having posed this question, people will provide answers that will change or nuance my conclusions.

REFERENCES

Ao, P. (2004). Potential in stochastic differential equations: novel construction. J. Phys. Math. Gen. *37*, L25–L30.

Bays, P.M., and Wolpert, D.M. (2007). Computational principles of sensorimotor control that minimize uncertainty and variability. J. Physiol. *578*, 387–396.

Bellman, R. (1952). On the Theory of Dynamic Programming. Proc. Natl. Acad. Sci. USA *38*, 716–719.

Bernstein, N.A. (1967). The Co-ordination and Regulation of Movements (Oxford: Pergamon Press).

Botvinick, M.M., and An, J. (2008). Goal-directed decision making in prefrontal cortex: a computational framework. In Advances in Neural Information Processing Systems (NIPS), D. Koller, Y.Y. Bengio, D. Schuurmans, L. Bouttou, and A. Culotta, eds. http://books.nips.cc/papers/files/nips21/NIPS2008_0034.pdf.

Braun, D., Ortega, P., Theodorou, E., and Schaal, S. (2011). Path Integral Control and Bounded Rationality. ADPRL 2011, (p. ID Code 7312). Paris.

Brown, L.D. (1981). A Complete Class Theorem for Statistical Problems with Finite Sample Spaces. Ann. Stat. *9*, 1289–1300.

Cooper, G. (1988). A Method for Using Belief Networks as Influence Diagrams. In Proceedings of the Conference on Uncertainty in Artificial Intelligence, pp. 55–63.

Dayan, P., and Hinton, G.E. (1997). Using expectation maximization for reinforcement learning. Neural Comput. *9*, 271–278.

Dayan, P., Hinton, G.E., Neal, R.M., and Zemel, R.S. (1995). The Helmholtz machine. Neural Comput. *7*, 889–904.

Diedrichsen, J., Shadmehr, R., and Ivry, R.B. (2010). The coordination of movement: optimal feedback control and beyond. Trends Cogn. Sci. (Regul. Ed.) *14*, 31–39.

Dvijotham, K., and Todorov, E. (2010). Inverse Optimal Control with Linearly-Solvable MDPs. Proceedings of the 27th International Conference on Machine Learning, pp. 335–342. Haifa.

Feldman, A.G. (2009). New insights into action-perception coupling. Exp. Brain Res. *194*, 39–58.

Feldman, A.G., and Levin, M.F. (1995). The origin and use of positional frames of reference in motor control. Behav. Brain Sci. *18*, 723–806.

Frens, M.A., and Donchin, O. (2009). Forward models and state estimation in compensatory eye movements. Front Cell Neurosci *3*, 13.

Friston, K. (2008). Hierarchical models in the brain. PLoS Comput. Biol. *4*, e1000211.

Friston, K. (2010). The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. *11*, 127–138.

Friston, K., and Ao, P. (2011). Free-energy, value and attractors. Computational and mathematical methods in medicine, in press.

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

Friston, K.J., Daunizeau, J., and Kiebel, S.J. (2009). Reinforcement learning or active inference? PLoS ONE *4*, e6421.

Friston, K.J., Daunizeau, J., Kilner, J., and Kiebel, S.J. (2010). Action and behavior: a free-energy formulation. Biol. Cybern. *102*, 227–260.

Friston, K., Mattout, J., and Kilner, J. (2011). Action understanding and active inference. Biol. Cybern. *104*, 137–160.

Gigerenzer, G., and Gaissmaier, W. (2011). Heuristic decision making. Annu. Rev. Psychol. *62*, 451–482.

Gläscher, J., Daw, N., Dayan, P., and O'Doherty, J.P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron *66*, 585–595.

Grafton, S.T., and Hamilton, A.F. (2007). Evidence for a distributed hierarchy of action representation in the brain. Hum. Mov. Sci. *26*, 590–616.

Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. Neuron *69*, 407–422.

Hoffman, M., de Freitas, N., Doucet, A., and Peters, J. (2009). An expectation maximization algorithm for continuous Markov decision processes with arbitrary rewards. Twelfth Int. Conf. on Artificial Intelligence and Statistics (AISTATS 2009), pp. 232–239.

Jensen, F., Jensen, V., and Dittmer, S.L. (1994). From influence diagrams to junction trees. In Proc. of the Tenth Conf. on Uncertainty in Artificial Intelligence (San Francisco: Morgan Kaufmann), pp. 367–373.

Jirsa, V.K., and Kelso, J.A. (2005). The excitator as a minimal model for the coordination dynamics of discrete and rhythmic movement generation. J. Mot. Behav. *37*, 35–51.

Kakei, S., Hoffman, D.S., and Strick, P.L. (2003). Sensorimotor transformations in cortical motor areas. Neurosci. Res. *46*, 1–10.

Kameswaran, S., and Biegler, L.T. (2006). Simultaneous dynamic optimization strategies: Recent advances and challenges. Computers and Chemical Engineering. *30*, 1560–1575.

Kappen, H.J. (2005). Linear theory for control of nonlinear stochastic systems. Phys. Rev. Lett. *95*, 200201.

Kappen, H.J., Gomez, Y., and Opper, M. (2009). Optimal control as a graphical model inference problem. arXiv, arXiv:0901.0633v2, http://arxiv.org/abs/0901.0633v2.

Kiebel, S.J., Daunizeau, J., and Friston, K.J. (2008). A hierarchy of time-scales and the brain. PLoS Comput. Biol. *4*, e1000209.

Körding, K. (2007). Decision theory: what "should" the nervous system do? Science *318*, 606–610.

Körding, K.P., and Wolpert, D.M. (2004). Bayesian integration in sensorimotor learning. Nature *427*, 244–247.

Littman, M.L., Majercik, S.M., and Pitassi, T. (2001). Stochastic boolean satisfiability. J. Autom. Reason. *27*, 251–296.

Liu, D., and Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. J. Neurosci. *27*, 9354–9368.

Miall, R.C. (2003). Connecting mirror neurons and forward models. Neuroreport *14*, 2135–2137.

Miall, R.C., Weir, D.J., Wolpert, D.M., and Stein, J.F. (1993). Is the cerebellum a smith predictor? J. Mot. Behav. *25*, 203–216.

Mitter, S., and Newton, N. (2003). A variational approach to nonlinear estimation. SIAM J. Contr. Optim. *42*, 1813–1833.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. Biol. Cybern. *66*, 241–251.

Mussa Ivaldi, F.A., Morasso, P., and Zaccaria, R. (1988). Kinematic networks. A distributed model for representing and regularizing motor redundancy. Biol. Cybern. *60*, 1–16.

Neal, R.M., and Hinton, G.E. (1998). A view of the EM algorithm that justifies incremental sparse and other variants. In Learning in Graphical Models, M. Jordan, ed. (Dordrecht: Kluwer Academic), pp. 355–368.

Paulin, M.G. (2005). Evolution of the cerebellum as a neuronal machine for Bayesian state estimation. J. Neural Eng. *2*, S219–S234.

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

Poeppel, D., Idsardi, W.J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. Philos. Trans. R. Soc. Lond. B Biol. Sci. *363*, 1071–1086.

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

Rawlik, K., Toussaint, M., and Vijayakumar, S. (2010). Approximate inference and stochastic optimal control. arXiv, arXiv:1009.3958, http://arxiv.org/abs/1009.3958.

Robert, C. (1992). L'analyse Statistique Bayesienne (Paris: Economica).

Schaal, S., Mohajerian, P., and Ijspeert, A. (2007). Dynamics systems vs. optimal control—a unifying view. Prog. Brain Res. *165*, 425–445.

Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. Annu. Rev. Neurosci. *23*, 473–500.

Shachter, R.D. (1988). Probabilistic inference and influence diagrams. Oper. Res. *36*, 589–605.

Shadmehr, R., and Krakauer, J.W. (2008). A computational neuroanatomy for motor control. Exp. Brain Res. *185*, 359–381.

Sherman, S.M., and Guillery, R.W. (1998). On the actions that one nerve cell can have on another: distinguishing "drivers" from "modulators". Proc. Natl. Acad. Sci. USA *95*, 7121–7126.

Shipp, S. (2005). The importance of being agranular: a comparative account of visual and motor cortex. Philos. Trans. R. Soc. Lond. B Biol. Sci. *360*, 797–814.

Sutton, R.S., and Barto, A.G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. Psychol. Rev. *88*, 135–170.

Tani, J. (2003). Learning to generate articulated behavior through the bottom-up and the top-down interaction processes. Neural Netw. *16*, 11–23.

Tani, J., Ito, M., and Sugita, Y. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using RNNPB. Neural Netw. *17*, 1273–1289.

Tassa, Y., Erez, T., and Todorov, E. (2011). Optimal limit-cycle control recast as Bayesian inference. World Congress of the International Federation of Automatic Control, pp. 4707–4713.

Theodorou, E., Buchli, J., and Schaal, S. (2010). A Generalized Path Integral Control Approach to Reinforcement Learning. J. Mach. Learn. Res. *11*, 3137–3181.

Todorov, E. (2004). Optimality principles in sensorimotor control. Nat. Neurosci. *7*, 907–915.

Todorov, E. (2008). General duality between optimal control and estimation. IEEE Conference on Decisionand Control, pp. 4286–4292.

Todorov, E., and Jordan, M.I. (2002). Optimal feedback control as a theory of motor coordination. Nat. Neurosci. *5*, 1226–1235.

Toussaint, M., and Storkey, A. (2006). Probabilistic inference for solving discrete and continuous state Markov Decision Processes. Proc. of the 23nd Int. Conf. on Machine Learning, pp. 945–952.

Toussaint, M., Charlin, L., and Poupart, P. (2008). Hierarchical POMDP controller optimization by likelihood maximization. Uncertainty in Artificial Intelligence (UAI 2008), pp. 562–570. AUAI Press.

Tseng, Y.W., Diedrichsen, J., Krakauer, J.W., Shadmehr, R., and Bastian, A.J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. J. Neurophysiol. *98*, 54–62.

van den Broek, B., Wiegerinck, W., and Kappen, B. (2008). Graphical model inference in optimal control of stochastic multi-agent systems. J. Artif. Intell. Res. *32*, 95–122.

Verschure, P.F., Voegtlin, T., and Douglas, R.J. (2003). Environmentally mediated synergy between perception and behaviour in mobile robots. Nature *425*, 620–624.

Wang, J.M., Fleet, D.J., and Hertzmann, A. (2009). Optimizing walking controllers. ACM Transactions on Graphics (TOG), SIGGRAPH Asia 2009, Article 168, pp. 8.

Wei, K., and Körding, K.P. (2009). Relevance of error: what drives motor adaptation? J. Neurophysiol. *101*, 655–664.

Wolpert, D.M., and Miall, R.C. (1996). Forward Models for Physiological Motor Control. Neural Netw. *9*, 1265–1279.

Wolpert, D.M., Ghahramani, Z., and Jordan, M.I. (1995). An internal model for sensorimotor integration. Science *269*, 1880–1882.

Wolpert, D.M., Diedrichsen, J., and Flanagan, J.R. (2011). Principles of sensorimotor learning. Nat Rev Neurosci., in press.

Wörgötter, F., and Porr, B. (2005). Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. Neural Comput. *17*, 245–319.

Zhang, N.L. (1998). Probabilistic inference in influence diagrams. Comput. Intell. *14*, 475–497.