Flexible Intentions in the Posterior Parietal Cortex: An Active Inference Theory

Matteo Priorelli[1] **and Ivilin Peev Stoianov**[1*]

***For correspondence:** ivilinpeev.stoianov@cnr.it (IPS)

[1]Institute of Cognitive Sciences and Technologies (ISTC), National Research Council of Italy (CNR), Padova, Italy

Abstract We present a normative computational theory of how neural circuitry in the Dorsal Visual Stream (DVS) and the Posterior Parietal Cortex (PPC) support visually guided goal-directed actions in a dynamical environment through flexible goal-encoding intentions. It builds on Active Inference, in which perception and motor control signals are inferred through dynamic minimization of generalized prediction error. The PPC is proposed to maintain expectations, or beliefs over environment and body state, and by manipulating them it is involved along with visual, frontal, and motor areas in dynamically generating goal-directed actions. Specifically, the PPC is viewed as a dynamic system computing flexible motor intentions and directing belief over the latent body state towards the dynamic future goals, while the DVS and proprioceptive pathways implement generative models translating belief into sensory-level predictions in order to infer targets, posture, and motor commands. A proof-of-concept agent embodying an actuated upper limb, visual and proprioceptive sensors and controlled by Active Inference extended with flexible intentions was tested on target reaching tasks. The agent behaved correctly under various conditions, including static and dynamic targets, different sensory feedbacks, sensory precisions, intention gains, and movement policies; limit conditions were individuated, too. Active Inference driven by flexible intentions can thus support goal-directed behavior in dynamically changing environments and the PPC can host its core intention mechanism. More broadly, the study provides normative basis for research on goal-directed behavior in end-to-end settings and further advances Artificial Intelligence methods for sensorimotor control.

INTRODUCTION

In primates, the dorsomedial visual stream provides critical support for continuously monitoring the spatial location of objects and posture to specify and guide actions, and performs visuomotor transformations in the course of evolving actions (***Fattori et al., 2017***; ***Galletti and Fattori, 2018***; ***Cisek and Kalaska, 2010***). The PPC, located at the apex of the dorsal stream, is also bidirectionally connected to frontal areas, motor and somatosensory cortex, placing it in a privileged position to set goal-directed actions and continuously adjusting motor plans by tracking moving targets and body posture (***Andersen, 1995***; ***Gamberini et al., 2021***) in a common reference frame (***Cohen and Andersen, 2002***). Undoubtedly, the PPC plays a crucial role in visually guided motor control (***Filippini et al., 2018***; ***Gamberini et al., 2021***; ***Desmurget et al., 1999***), but its specific role and mechanism are still disputed. The most consistent view is that the PPC estimates the state of the body and the environment and optimizes the body-environment interactions (***Medendorp and Heed, 2019***). Others see the PPC as a task estimator (***Haar and Donchin, 2020***), and even more classically, that it is

involved in endogenous attention and task setting (***Corbetta and Shulman, 2002***). Its underlying computational mechanism is not fully understood, especially as regards the definition of goals in motor planning and their integration within the control process (***Shadmehr and Krakauer, 2008***). For example, the prevailing Optimal Feedback Control theory defines motor goals through task-specific cost functions (***Todorov, 2004***). Neural-level details of motor-goal coding are becoming increasingly more important in light of the growing demand for neural interfaces targeting brain areas providing information about the motor intent (***Gallego et al., 2022***) in support of intelligent assistive devices (***Velliste et al., 2008***; ***Srinivasan et al., 2021***).

Intentions encode motor goals set before the beginning of motor acts themselves and are therefore viewed as memory holders of voluntary actions (***Andersen, 1995***; ***Snyder et al., 1997***; ***Fogassi et al., 2005***; ***Lau et al., 2004***). Several cortical areas handle different aspects of intentions: the premotor cortex (PM) encodes structuring while the Supplementary Motor Area (SMA) controls phasing (***Gallego et al., 2022***). In turn, PPC plays a role in dynamic action tuning, different PPC neurons being sensitive to different intentions (***Snyder et al., 2000***). Notably, intention neurons respond not only when performing a given action but also during its observation, allowing observers to «predict the goal of the observed action and, thus, to "read" the intention of the acting individual» (***Fogassi et al., 2005***). Motor goals have also been observed down the motor hierarchy, which is an expression of hierarchical predictive coding in the motor domain (***Friston et al., 2011***).

To investigate how the PPC supports sensory guided actions through motor intentions from a computational point of view, we adopted the Active Inference theory of cognitive and motor control of behavior, which provides with increasing appeal fundamental insights about the computational role and principles of the nervous system, especially with regard to the perception-action loop (***Friston and Kiebel, 2009***; ***Friston et al., 2010***; ***Bogacz, 2017***; ***Parr et al., 2021***). Active Inference provides a formalization of two main computational tasks of the cortex, namely the interpretation of sensory information, i.e., perception and control, both of which are viewed as aiming to resolve the critical goal of the organisms: to survive in uncertain environments by operating within preferred states (e.g., maintaining a constant temperature). Accordingly, both these tasks are implemented by dynamic minimization of a quantity called *free energy*, whose process generally corresponds to the minimization of high- and low-level prediction errors, that is, satisfaction of prior and sensory expectations. There are two branches of Active Inference that are appropriate to tackle two different levels of control. The discrete branch can explain high-level cognitive control processes such as planning and decision making; it evaluates expected outcomes to select actions in discrete entities (***Pezzulo et al., 2018***). In turn, continuous adjustment of action plans in the PPC considering the visual stream matches by functionality dynamic Active Inference in continuous state space, which formalizes perception and sensorimotor control by casting them in terms of reduction of prediction errors at both internal and sensory levels (***Friston et al., 2011***, ***2012***). Therefore, this framework departs from the classical view of perception, planning, and motor control, considering them as a dynamic probabilistic inference problem (***Kaplan and Friston, 2018***; ***Toussaint and Storkey, 2006***; ***Millidge et al., 2020***; ***Levine, 2018***). The biologically implausible cost functions typical of optimal control are replaced by high-level priors defined in extrinsic state-space, allowing complex movements such as handwriting (***Friston, 2011***; ***Adams et al., 2013***).

In the following, we first outline the background computational framework and then elaborate on intentionality in dynamic Active Inference. Our most critical contributions regard the formalization of goal-directed behavior and the processes linking flexible goals (e.g., dynamic visual targets) with motor plans through the definition of intention functions. We also investigate a more parsimonious approach to motor control based solely on proprioceptive predictions. We then provide implementation details and a practical demonstration of the theoretical contribution in terms of a simulated Active Inference agent, which we show is capable of detecting and reaching static visual goals and tracking moving targets. We also provide detailed performance statistics and investigate the effects of key system parameters whose balance is critical to movement stability. Additionally, gradient analysis provides critical insight into the causes of the movements performed. Finally, we

discuss how intentions could be selected to perform a series of goal-directed steps, e.g., a multi-phase action, and illustrate conditions for neurological disorders.

COMPUTATIONAL BACKGROUND

We first outline the computational principles of the underlying probabilistic and predictive coding approach and provide background on variational inference, free energy minimization, Active Infefence, and variational autoencoders necessary to comprehend the following main contribution.

Bayesian Brain

In recent years, there has been increasing attention to a radically new theory of the mind called the *Bayesian brain*, according to which the brain is a sophisticated machine that constantly makes use of Bayesian reasoning to capture causal relationships in the world and deliver optimal behavior in an uncertain environment (***Hohwy, 2013***; ***Doya, 2007***; ***Pezzulo et al., 2017***). At the core of the theory is in fact the Bayes theorem, whose application here implies that posterior beliefs about the world are updated according to the product of prior beliefs and the likelihood of observing sensory input. Priors play a fundamental role in driving the dynamics of perceptual experience, as evident in visual phenomena such as illusions or binocular rivalry. Perception is then more of a simple bottom-up feedforward mechanism that detects features and objects from the sensorium and computes latent states; rather, it comprises a predictive top-down generative model which continuously anticipates the sensory input in order to test hypotheses and explain away ambiguities.

According to Bayesian brain theory, this apparently complex task is accomplished through *predictive coding*, implemented by means of message passing of top-down predictions and bottom-up prediction errors between adjacent layers in the cortical hierarchy (***Rao and Ballard, 1999***). The former are generated from latent states maintained at the highest levels, representing beliefs about the causes of the environment, while the latter are computed by comparing sensory-level predictions with the actual observations. Each prediction will then act as a cause for the layer below, while the prediction error will convey information to the layer above. It is thanks to this hierarchical organization and through error minimization at every layer that the cortex is able to mimic and capture the inherently hierarchical relationships that model the world. In this view, sensations are only needed in that they provide, through the computation of prediction errors, a measure of how good the model is and a cue to correct future predictions. Thus, ascending projections do not encode the features of a stimulus, but rather how much the brain is *surprised* about it, considering the strict correlation between surprise and model uncertainty.

Variational Bayes

Organisms are supposed to implement model fit or error minimization by some form of variational inference, a broad family of techniques based on the *calculus of variations* and used to approximate intractable posteriors that would otherwise be infeasible to compute analytically or even with classical sampling methods such as Monte Carlo (***Bishop, 2006***).

Nervous systems compute control signals under uncertainty about the real, yet unknown, state of the world, including the state of the organisms per se, which could jointly be denoted by a latent variable $\boldsymbol{z}$. If organisms are endowed with sensors that provide partial evidence $p(\boldsymbol{s})$ on the state of the dynamic world and have prior knowledge $p(\boldsymbol{z})$, they could use sensory evidence and apply Bayesian inference to improve their knowledge (***Ma et al., 2006***). To do so, given the observation $\boldsymbol{s}$, the nervous system needs to evaluate the posterior $p(\boldsymbol{z}|\boldsymbol{s})$:

$$p(\boldsymbol{z}|\boldsymbol{s}) = \frac{p(\boldsymbol{z}, \boldsymbol{s})}{p(\boldsymbol{s})} \tag{1}$$

However, evaluating this posterior is infeasible due to the intractability of the marginal $p(\boldsymbol{s}) = \int p(\boldsymbol{z}, \boldsymbol{s}) d\boldsymbol{z}$, which involves integration over the joint density $p(\boldsymbol{z}, \boldsymbol{s})$.

The variational approach approximates the posterior with a simpler to compute *recognition* distribution $q(\boldsymbol{z}) \approx p(\boldsymbol{z}|\boldsymbol{s})$, assuming furthermore that a neural process, e.g., perceptual learning, makes those distributions similar by minimizing the Kullback-Leibler (KL) divergence between them:

$$D_{KL}[q(\boldsymbol{z})||p(\boldsymbol{z}|\boldsymbol{s})] = \int_{\boldsymbol{z}} q(\boldsymbol{z}) \ln \frac{q(\boldsymbol{z})}{p(\boldsymbol{z}|\boldsymbol{s})} d\boldsymbol{z} \tag{2}$$

The KL divergence can be rewritten as the difference between log-evidence $\ln p(\boldsymbol{s})$ and a quantity $\mathcal{L}(q)$ known as *evidence lower bound*, or ELBO (***Bishop, 2006***):

$$D_{KL}[q(\boldsymbol{z})||p(\boldsymbol{z}|\boldsymbol{s})] = \ln p(\boldsymbol{s}) - \int_{\boldsymbol{z}} q(\boldsymbol{z}) \ln \frac{p(\boldsymbol{z}, \boldsymbol{s})}{q(\boldsymbol{z})} d\boldsymbol{z} = \ln p(\boldsymbol{s}) - \mathcal{L}(q) \tag{3}$$

Since the KL divergence is always nonnegative, the ELBO provides a lower bound on log-evidence, that is, $\mathcal{L}(q) \leq \ln p(\boldsymbol{s})$. Therefore, minimizing the KL divergence with respect to $q(\boldsymbol{z})$ is equivalent to maximizing $\mathcal{L}(q)$, which at its maximum corresponds to an approximate density that is closest the most to the posterior, depending on the particular choice of the form of $q(\boldsymbol{z})$. In general, very few assumptions are made about the form that this distribution can take (a multivariate Gaussian is a typical choice), with a trade-off between having a tractable optimization process while still leading to a good approximate posterior.

Free Energy

Why has the cortex evolved in such a way as to realize Bayesian inference? And how can it be implemented through simple message passing of prediction errors? Friston (***2002***; ***2005***) proposed an elegant solution based on the so-called *free energy*, a concept borrowed from thermodynamics and defined as the negative ELBO. Accordingly, Equation (3) can be rewritten as:

$$F(\boldsymbol{z}, \boldsymbol{s}) = -\mathcal{L}(q) = D_{KL}[q(\boldsymbol{z})||p(\boldsymbol{z}|\boldsymbol{s})] - \ln p(\boldsymbol{s}) = \int_{\boldsymbol{z}} q(\boldsymbol{z}) \ln \frac{q(\boldsymbol{z})}{p(\boldsymbol{z}, \boldsymbol{s})} d\boldsymbol{z} \tag{4}$$

Minimizing free energy is then equivalent to maximization of ELBO and provides an upper bound on surprise, a process called *perceptual inference*:

$$\boldsymbol{z} = \arg\min_{\boldsymbol{z}} F(\boldsymbol{z}, \boldsymbol{s}) \tag{5}$$

In this way, the organism indirectly minimizes surprise (hence model uncertainty) and is able to learn the causal relations between unknown states and sensory input, and generate predictions based on its current internal representation of the environment.

Free energy minimization is simpler than dealing with the KL divergence between the approximate and true posterior because the former depends on quantities that the organism has access to, namely the approximate posterior and the generative model.

Assuming (i) under the mean-field approximation that the recognition density can be partitioned into independent distributions: $q(\boldsymbol{z}) = \prod_i q(\boldsymbol{z}_i)$, and (ii) under the Laplace approximation that each of these partitions is Gaussian: $q(\boldsymbol{z}_i) = \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Pi}_i^{-1})$, where $\boldsymbol{\mu}_i$ represents the most plausible hypothesis (also called *belief* about the hidden state $\boldsymbol{z}_i$), and $\boldsymbol{\Pi}_i$ is its precision matrix that could be treated as a parameter or learned online (***Friston et al., 2007***), the free energy becomes independent of $\boldsymbol{z}$ and simplifies as follows:

$$F(\boldsymbol{\mu}, \boldsymbol{s}) = -\ln p(\boldsymbol{\mu}, \boldsymbol{s}) + C = -\ln p(\boldsymbol{s}|\boldsymbol{\mu}) - \ln p(\boldsymbol{\mu}) + C \tag{6}$$

The dynamic environmental process can be modeled with the following nonlinear stochastic system:

$$\begin{aligned} \boldsymbol{s} &= \boldsymbol{g}(\boldsymbol{z}) + \boldsymbol{w}_s \\ \dot{\boldsymbol{z}} &= \boldsymbol{f}(\boldsymbol{z}) + \boldsymbol{w}_z \end{aligned} \tag{7}$$

where $\boldsymbol{g}$ is a function that converts latent states or causes $\boldsymbol{z}$ into observed states or sensations $\boldsymbol{s}$, and the function $\boldsymbol{f}$ encodes the dynamics of the system, that is, the change in latent states

$\dot{z}$. In turn, $\boldsymbol{w}_s$ and $\boldsymbol{w}_z$ are noise terms that describe system uncertainty, assumed to belong to multivariate normal distributions with zero mean and respective precisions $\boldsymbol{\Pi}_s$ and $\boldsymbol{\Pi}_z$.

As explained above, under the variational approach, one can assume that $\boldsymbol{z}$ is drawn from a Gaussian distribution with mean $\boldsymbol{\mu}$ so that the functions $\boldsymbol{g}$ and $\boldsymbol{f}$ correspond to the likelihood $p(\boldsymbol{s}|\boldsymbol{\mu})$ and the density $p(\boldsymbol{\mu}'|\boldsymbol{\mu})$ respectively, where $\boldsymbol{\mu}'$ is a belief over the 1st-order dynamics. We can have a more precise description of the unknown environmental dynamics by considering not only the 1st-order of Equation (7), but also higher temporal orders, using respective approximations of the higher order states: $\tilde{\boldsymbol{\mu}} = \{\boldsymbol{\mu}, \boldsymbol{\mu}', \boldsymbol{\mu}'', \ldots\}$ - called *generalized coordinates* (***Friston, 2008***; ***Friston et al., 2008***). This allows to better represent the environmental dynamics with the following generalized model:

$$\begin{aligned} \tilde{\boldsymbol{s}} &= \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{\mu}}) + \boldsymbol{w}_s \\ D\tilde{\boldsymbol{\mu}} &= \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{\mu}}) + \boldsymbol{w}_\mu \end{aligned} \tag{8}$$

where D is the differential (shift) operator matrix such that $D\tilde{\boldsymbol{\mu}} = \{\boldsymbol{\mu}', \boldsymbol{\mu}'', \ldots\}$, $\tilde{\boldsymbol{s}}$ denotes the generalized sensors, while $\tilde{\boldsymbol{g}}$ and $\tilde{\boldsymbol{f}}$ denote the generalized model functions for each temporal order. Note that in this system, the sensory data at a particular dynamical order $\boldsymbol{s}^{[d]}$ - where $[d]$ is the order - engage only with the same order of belief $\boldsymbol{\mu}^{[d]}$, while the generalized equation of motion, or system dynamics, specifies the coupling between adjacent orders. We can approximate the generalized environmental model with joint densities of all dynamical orders and expand them as follows:

$$\begin{aligned} p(\tilde{\boldsymbol{s}}|\tilde{\boldsymbol{\mu}}) &= \prod_d p(\boldsymbol{s}^{[d]}|\boldsymbol{\mu}^{[d]}) \\ p(\tilde{\boldsymbol{\mu}}) &= \prod_d p(\boldsymbol{\mu}^{[d+1]}|\boldsymbol{\mu}^{[d]}) \end{aligned} \tag{9}$$

As defined above, these variational probability distributions are assumed to be Gaussian:

$$\begin{aligned} p(\boldsymbol{s}^{[d]}|\boldsymbol{\mu}^{[d]}) &= \frac{\boldsymbol{\Pi}_s}{\sqrt{(2\boldsymbol{\pi})^L}} \exp\left(-\frac{1}{2}\boldsymbol{e}_s^{[d]T}\boldsymbol{\Pi}_s\boldsymbol{e}_s^{[d]}\right) \\ p(\boldsymbol{\mu}^{[d+1]}|\boldsymbol{\mu}^{[d]}) &= \frac{\boldsymbol{\Pi}_\mu}{\sqrt{(2\boldsymbol{\pi})^M}} \exp\left(-\frac{1}{2}\boldsymbol{e}_\mu^{[d]T}\boldsymbol{\Pi}_\mu\boldsymbol{e}_\mu^{[d]}\right) \end{aligned} \tag{10}$$

where L and M are the dimensions of sensations and belief, respectively. Note that the probability distributions are expressed in terms of sensory and belief dynamics prediction errors:

$$\boldsymbol{e}_s^{[d]} = \boldsymbol{s}^{[d]} - \boldsymbol{g}^{[d]}(\boldsymbol{\mu}^{[d]}) \tag{11}$$

$$\boldsymbol{e}_\mu^{[d]} = \boldsymbol{\mu}^{[d+1]} - \boldsymbol{f}^{[d]}(\boldsymbol{\mu}^{[d]}) \tag{12}$$

The factorized probabilistic approximation of the dynamic model allows easy generalized state (belief) estimation based on free energy minimization performed by iterative gradient descent over generalized coordinates, that is, by changing the belief $\tilde{\boldsymbol{\mu}}$ about all orders of the hidden states:

$$\dot{\tilde{\boldsymbol{\mu}}} - D\tilde{\boldsymbol{\mu}} = -\partial_{\tilde{\boldsymbol{\mu}}} F(\tilde{\boldsymbol{s}}, \tilde{\boldsymbol{\mu}}) \tag{13}$$

Gradient descent is tractable because the Gaussian variational functions are smooth and differentiable and the derivatives are easily computed in terms of generalized prediction errors, since the logarithm of Equation (6) vanishes the exponent of the Gaussian. The belief update thus turns to:

$$\dot{\tilde{\boldsymbol{\mu}}} = D\tilde{\boldsymbol{\mu}} + \frac{\partial\tilde{\boldsymbol{g}}}{\partial\tilde{\boldsymbol{\mu}}}^T \tilde{\boldsymbol{\Pi}}_s\tilde{\boldsymbol{e}}_s + \frac{\partial\tilde{\boldsymbol{f}}}{\partial\tilde{\boldsymbol{\mu}}}^T \tilde{\boldsymbol{\Pi}}_\mu\tilde{\boldsymbol{e}}_\mu - D^T\tilde{\boldsymbol{\Pi}}_\mu\tilde{\boldsymbol{e}}_\mu \tag{14}$$

which is made up of three main terms: a likelihood error computed at the sensory-level, a backward error coming from the next temporal order, and a forward error from the previous order. These terms represent the free energy gradients relative to the belief terms $\boldsymbol{\mu}^{[d]}$ of Equation (11), and $\boldsymbol{\mu}^{[d+1]}$ and $\boldsymbol{\mu}^{[d]}$ of Equation (12), respectively.

In conclusion, by making some plausible simplifying assumptions, the complexity of free energy minimization has been reduced to the computation of predictions, which are constantly compared with the sensory observations so to determine a prediction error signal. This error then flows back through the cortical hierarchy to adjust the distribution parameters accordingly and minimize sensory surprise, which is equivalent to maximizing evidence, in the long run.

Active Inference

Learning the causal relationships of some observed data (e.g., what causes an increase in body temperature) is not sufficient to keep organisms alive (e.g., maintaining temperature in a vital range). Critically, an organism can also modify sensory states by interacting with the environment. It has been noted that the Bayesian brain hypothesis can be seen as a corollary of a more general *free energy principle* (FEP), which states that, in order for an organism to maintain a state of homeostasis and survive, it must constantly and actively restrict the set of latent states in which it lives to a narrow range of life-compatible possibilities, counteracting the natural tendency for disorder (***Friston, 2012***) - hence the relationship with thermodynamics. If these states are defined by the phenotype, from the point of view of the organism's internal model they are exactly the states that it expects to be less surprising. Thus, while perceptual inference tries to optimize the belief about hidden causes in order to explain away sensations, on the other hand, if the assumptions defined by the phenotype are considered to be the true causes of the world, by interacting with the external environment the agent will try to sample those sensations that make the assumptions true, fulfilling its needs and beliefs: *active inference* then becomes a self-fulfilling prophecy. In this view, there is no difference between a desire or a belief: we simply seek the states in which we expect to find ourselves (***Friston et al., 2010***; ***Buckley et al., 2017***).

To perform goal-directed behavior, it is then sufficient to apply free energy minimization also with respect to the action (see Equation 6):

$$\boldsymbol{a} = \arg\min_{\boldsymbol{a}} F(\boldsymbol{\mu}, \boldsymbol{s}) \tag{15}$$

Given that it depends only on the sensory information, we obtain:

$$\dot{\boldsymbol{a}} = -\partial_{\boldsymbol{a}} F(\tilde{\boldsymbol{s}}, \tilde{\boldsymbol{\mu}}) = -\frac{\partial F}{\partial \tilde{\boldsymbol{s}}} \frac{\partial \tilde{\boldsymbol{s}}}{\partial \boldsymbol{a}} = -\frac{\partial \tilde{\boldsymbol{s}}}{\partial \boldsymbol{a}}^T \tilde{\boldsymbol{\Pi}}_s \tilde{\boldsymbol{e}}_s \tag{16}$$

Computing the action by minimizing the free energy of all sensations is certainly useful, since it makes use of every contribution of the likelihood that drives the belief update; however, it requires the knowledge of an inverse mapping from exteroceptive sensations to actions (***Baltieri and Buckley, 2019***), which is generally considered a "hard problem" because the mapping is highly non-linear and not univocal (***Friston et al., 2010***).

In a more realistic scenario, only proprioception drives the minimization of free energy with respect to action, which is simpler to compute since the corresponding sensory information is already in the intrinsic proprioceptive domain. Control signals sent from the motor cortex, therefore, are not motor commands as in the classical view of optimal control; instead, they consist of predictions that specify the desired trajectory. In this scenario, proprioceptive prediction errors computed locally at the spinal cord serve two purposes that only differ in how these signals are suppressed: they drive the current belief towards sensory observations (happening to realize perception, as with exteroceptive signals) and they drive sensory observations towards the current belief (happening to realize movement), further suppressing the same errors by means of simple reflex arcs that activate the corresponding muscles (***Adams et al., 2013***; ***Parr and Friston, 2018***; ***Versteeg et al., 2021***).

In conclusion, perception and action can be seen as two sides of the same coin implementing the common vital goal of minimizing entropy or average surprise. In this view, what we perceive never tries to perfectly match the real state of affairs of the world, but is constantly biased towards our preferred states. In turn, action only indirectly fulfills future goals; instead, it continuously tries to fill the gap between sensations and predictions generated from our biased belief.

Variational Autoencoder

Variational Autoencoders (VAEs) belong to the family of generative models, since they learn the joint distribution $p(\boldsymbol{z}, \boldsymbol{s})$ and can generate synthetic data similar to the input, given a prior distribution $p(\boldsymbol{z})$ over the latent space. VAEs use the variational Bayes approach to capture the posterior distribution $p(\boldsymbol{z}|\boldsymbol{s})$ of the latent representation of the inputs when the computation of the marginal is intractable (***Goodfellow et al., 2016***). A VAE is composed of two probability distributions, both of which are assumed to be Gaussian: a probabilistic *encoder* corresponding to the recognition distribution $q(\boldsymbol{z}|\boldsymbol{s})$ producing a distribution over the latent space $\boldsymbol{z}$ given input $\boldsymbol{s}$, and a generative function $p(\boldsymbol{s}|\boldsymbol{z})$ called probabilistic *decoder* which computes a distribution over the input space given a latent representation $\boldsymbol{z}$ (Figure 3c):

$$\begin{aligned} q(\boldsymbol{z}|\boldsymbol{s}) &= \mathcal{N}(\boldsymbol{z}|\boldsymbol{\mu}_\phi, \boldsymbol{\Sigma}_\phi) \\ p(\boldsymbol{s}|\boldsymbol{z}) &= \mathcal{N}(\boldsymbol{s}|\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) \end{aligned} \tag{17}$$

Although VAEs have many similarities with traditional autoencoders, they are actually a derivation of the AEVB algorithm when a neural network is used for the recognition distribution (***Kingma and Welling, 2014***).

Unlike other variational techniques, the approximate posterior is generally not assumed to be factorial, but since the calculation of $\nabla_\phi \mathcal{L}_{\theta,\phi}(\boldsymbol{s})$ is biased, a method called *reparametrization trick* is used to rewrite the gradient so that it is independent of the parameters $\boldsymbol{\phi}$.

This method works by expressing the latent variable $\boldsymbol{z}$ by a function:

$$\boldsymbol{z} = r(\epsilon, \boldsymbol{\phi}, \boldsymbol{s}) \tag{18}$$

where ϵ is an auxiliary variable independent of $\boldsymbol{\phi}$ and $\boldsymbol{s}$. The ELBO for a single datapoint can thus be expressed as:

$$\tilde{\mathcal{L}}_{\theta,\phi}(\boldsymbol{s}) = -D_{KL}[q(\boldsymbol{z}|\boldsymbol{s})||p(\boldsymbol{z})] + \frac{1}{L}\sum^{L} \log p(\boldsymbol{s}|\boldsymbol{z}_l) \tag{19}$$

which can be minimized through backpropagation. Here, the KL divergence can be seen as a regularizer, and the second RHS term is an expected negative reconstruction term.

RESULT

In what follows, we develop and preliminary assess a computational theory of the circuitry in the PPC and related areas controlling goal-directed actions in a dynamically changing environment through flexible intentions. We first elaborate on intentionality in Active Inference, then provide a proof-of-concept agent endowed with visual input, and finally describe and analyze several simulations of visually guided behavior. The theoretical work is motivated by basic research showing the critical role of PPC in goal-directed sensorimotor control through intention coding (***Andersen, 1995***; ***Desmurget et al., 1999***; ***Galletti and Fattori, 2018***) and extends previous theoretical and applied research on Active Inference (***Friston et al., 2009***; ***Pio-Lopez et al., 2016***; ***Limanowski and Friston, 2020***) and VAE-based vision support (***Lanillos and Cheng, 2018***; ***Sancaktar et al., 2020***; ***Rood et al., 2020***). In turn, the simulations were inspired by a monkey reaching task (***Breveglieri et al., 2014***).

Intentions

State-of-the-art implementations of dynamic Active Inference define action goals in terms of an attractor embedded in the system dynamics (***Friston, 2011***). The attractor can be a low-level (sensory) prediction that is compared with the current observation and backpropagated to obtain a belief update direction (***Sancaktar et al., 2020***; ***Oliver et al., 2019***; ***Pio-Lopez et al., 2016***), or a desired latent state specified in the full equations of motion that the agent expects to perceive (***Baioumy et al., 2020***). However, since these states or sensations are usually static or defined externally, the agent is not able to deal with continuously changing environments (e.g., reaching or grasping moving targets) expecting that the world will always evolve in the same way.

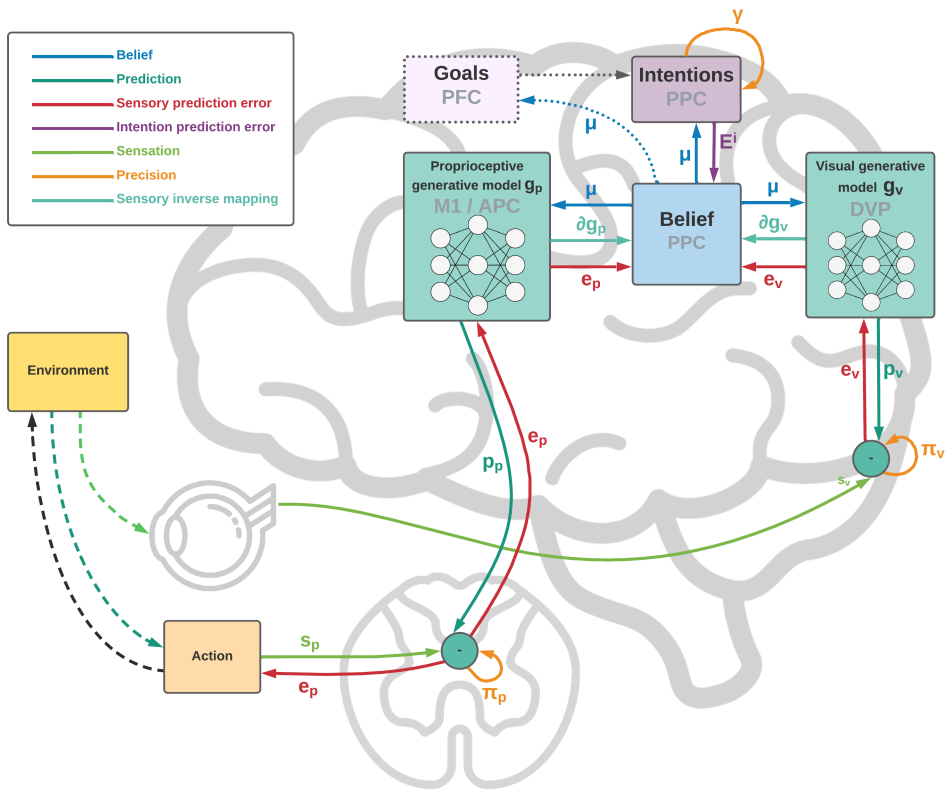


Figure 1. Functional architecture and cortical overlay of the goal-directed Active Inference agent. The process starts with the computation of future intentions $\boldsymbol{h}$ in the PPC under the coordination of frontal and motor areas. In the middle of the sensorimotor hierarchy, at the PPC, the dynamic Active Inference process maintains beliefs $\boldsymbol{\mu}$ over the latent causes of the current sensory observations $\boldsymbol{s}_p$ and $\boldsymbol{s}_v$, and computes proprioceptive and visual predictions through the somatosensory and dorsal visual pathways (for simplicity we have omitted the somatomotor pathway and considered a single mechanism for both motor control and belief inference). In turn, the lower layers of the hierarchy compute sensory prediction errors $\boldsymbol{e}_p$ and $\boldsymbol{e}_v$ while the higher layers compute intention prediction errors $\boldsymbol{e}^{(i)}$; both are propagated back towards the PPC, which thus integrates information from multiple sensory modalities and intentions. Free energy is minimized throughout the cortical hierarchy by changing the belief about the causes of the current observation (perception) and by sending proprioceptive predictions from the motor cortex to the reflex arcs (action). An essential element of this process is the computation of gradients $\partial \boldsymbol{g}_p$ and $\partial \boldsymbol{g}_v$ by inverse mappings from the sensations towards the deepest latent states. In this process, intentions act as high-level attractors and the belief propagated down to compute sensorimotor predictions embeds a component directing the body state towards the goals.

Instead, the agent could maintain a dynamically estimated belief over moving objects in the scene - potential or current targets - or be given a desired latent state to reach (***Friston et al., 2010***); in the brain, goal-encoding signals could descend from higher-level cortical structures (e.g., PFC, (***Stoianov et al., 2016***)). An intention can be computed by exploiting this belief in order to compute a future action goal, or body posture, so that the attractor - specified either in the belief domain or at the sensory-level - is not fixed but depends on the current perceptual and internal representation of the world, or representations memorized in past experiences. In other words, the agent can construct and maintain at a high level a plausible dynamic hypothesis for the causes that produce what it is perceiving; by manipulating these and eventually other memorized causes, the agent can dynamically compute a representation of the future body state, i.e., intention, which in turn can act as a prior over the current belief.

As shown in Figure 1, the belief is located at an intermediate level between generative mod-

els that produce sensory predictions and a system dynamics model that specifies its evolution over time. We propose to decompose the belief dynamics model into a set of intention functions that flexibly generate future beliefs, each intention providing an independent expectation that the agent will find itself in a particular state (***Friston et al., 2011***). The belief is then constantly subject to two forces: one from below (corresponding to *sensory prediction errors*) that pulls the belief to what the agent is currently perceiving, and one from higher hierarchical levels (which we can call *intention prediction errors*) that pulls the belief to what the agent expects to perceive in the future.

From a neural perspective, the PPC is the ideal candidate for a cortical structure computing beliefs and flexible intentions, being at the apex of the DVS hierarchy and other exteroceptive domains producing sensory predictions on the one hand, and on the other, linked with motor and frontal areas producing continuous trajectories in the proprioceptive domain and higher-level plans composed of a series of discrete action chunks. In fact, it is known that the PPC is an associative region that integrates information from multiple sensory modalities and encodes visuomotor transformations; for example, area V6A encodes object affordances during reaching and grasping tasks (***Filippini et al., 2017***; ***Fattori et al., 2017***). Thus, if the job of sensory pathways, or generative models, is to compute sensory-level predictions, our hypothesis is that higher levels of the sensorimotor control hierarchy integrate in the PPC previous states of belief with flexible priors computing (high-level) intentions.

To begin describing a more formal definition, we assume that the neural system perceives the environment and receives feedback on the effects of motor control through L noisy sensors $\boldsymbol{s}$. Here, $\boldsymbol{s}$ encloses sensors from multiple domains, most critically, proprioceptive and visual. Under the VB and Gaussian approximation of the recognition density, we also assume that the nervous system operates on beliefs $\boldsymbol{\mu} \in \mathbb{R}^M$ that define an abstract internal representation of the world. Furthermore, we assume that the agent maintains generalized coordinates up to the 2nd-order resulting from free energy minimization in the generalized belief space $\tilde{\boldsymbol{\mu}} = \{\boldsymbol{\mu}, \boldsymbol{\mu}', \boldsymbol{\mu}''\}$.

We then define *intentions* $\boldsymbol{h}$ as predictions of target goal states of the world over current beliefs $\boldsymbol{\mu}$ computed with the help of K *intention functions* $\boldsymbol{i}^{(i)}(\boldsymbol{\mu}) \in \mathbb{R}^{\mathbb{M}}$ building intention representations $\boldsymbol{h}^{(i)}$ for each goal of the agent. Although both the belief and the intentions could be very abstract representations of the world (comprising states in extrinsic and intrinsic coordinates), we assume a simpler scenario in which the intentions operate on beliefs in a common intrinsic motor-related domain, e.g., the joint angle space. We also assume that there are two conceptually different components in both the belief (domain) $\boldsymbol{\mu}$ and the output (codomain) of the intention functions $\boldsymbol{i}^{(i)}$. The first component could represent the bodily states and serve to drive actions, while the second one could represent the state of other objects, most importantly targets to interact with, which can be internally encoded in the joint angle space as well. These targets could be observed, but they could also be imagined or set by higher level cognitive control frontal areas such as PFC or PMd (***Genovesio et al., 2012***; ***Stoianov et al., 2016***).

For the sake of notational simplicity, we group all intentions into a single matrix $\boldsymbol{H} \in \mathbb{R}^{MxK}$:

$$\boldsymbol{H} = \boldsymbol{i}(\boldsymbol{\mu}) = \begin{bmatrix} \boldsymbol{i}^{(0)}(\boldsymbol{\mu}) & \dots & \boldsymbol{i}^{(K)}(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}^{(0)} & \dots & \boldsymbol{h}^{(K)} \end{bmatrix} \tag{20}$$

Intention prediction errors $\boldsymbol{e}^{(i)}$ are then defined as the difference between the current belief and every intention:

$$\boldsymbol{E}^{(i)} = \boldsymbol{H} - \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{e}^{(0)} & \dots & \boldsymbol{e}^{(K)} \end{bmatrix} \tag{21}$$

In turn, sensory predictions are produced by a set of generative models $\boldsymbol{g}_j$, one for each sensory modality. We group the predictions into a prediction vector $\boldsymbol{p}$:

$$\boldsymbol{p} = \boldsymbol{g}(\boldsymbol{\mu}) = \begin{bmatrix} \boldsymbol{g}_0(\boldsymbol{\mu}) \\ \vdots \\ \boldsymbol{g}_L(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{p}_0 \\ \vdots \\ \boldsymbol{p}_L \end{bmatrix} \tag{22}$$

Note that each term $\boldsymbol{p}_j$ is a multidimensional sensory-level representation that provides predictions for a particular sensory domain, with its own dimensionality, and grouping, for notational simplicity, all predicted bits from this domain into a single term.

Sensory prediction errors $\boldsymbol{e}_j$ are then computed as the difference between sensations from each domain and the corresponding sensory-level predictions:

$$\boldsymbol{e}_s = \boldsymbol{s} - \boldsymbol{p} = \begin{bmatrix} \boldsymbol{e}_0 \\ \vdots \\ \boldsymbol{e}_L \end{bmatrix} \tag{23}$$

where $\boldsymbol{s}$ encompasses the current sensory observations of all sensory modality, most importantly visual and proprioceptive. In the above notation and hereafter, the superscripts index intentions, and the subscripts index sensory domains.

Goal-directed Active Inference

Under the assumption of independence among intentions and sensations, we can factorize the joint probability of the generative model into a product of distributions for each sensory modality and intention, which expands as follows:

$$p(\boldsymbol{\mu}, \boldsymbol{s}) = \prod_i^K p(\boldsymbol{\mu}^{(i)\prime}|\boldsymbol{\mu}) \cdot \prod_j^L p(\boldsymbol{s}_j|\boldsymbol{\mu}) \tag{24}$$

The probability distributions are assumed to be Gaussian:

$$\begin{aligned} p(\boldsymbol{\mu}^{(i)\prime}|\boldsymbol{\mu}) &= \mathcal{N}(\boldsymbol{\mu}^{(i)\prime}|\boldsymbol{f}^{(i)}(\boldsymbol{\mu}), \gamma^{(i)}) \\ p(\boldsymbol{s}_j|\boldsymbol{\mu}) &= \mathcal{N}(\boldsymbol{s}_j|\boldsymbol{g}_j(\boldsymbol{\mu}), \pi_j) \end{aligned} \tag{25}$$

where $\gamma^{(i)}$ and π_j are, respectively, the precisions of intention i and sensor j. Here, $\boldsymbol{\mu}^{(i)\prime}$ and $\boldsymbol{f}^{(i)}$ correspond to the i-th component of the generalized dynamics function:

$$\boldsymbol{f}^{(i)}(\boldsymbol{\mu}) = k\boldsymbol{e}^{(i)} + \boldsymbol{w}_\mu^{(i)} \tag{26}$$

where k is the gain of intention prediction errors $\boldsymbol{E}^{(i)}$. Note that the goal states are embedded into the 1st-order dynamics functions, acting as belief-level attractors for each intention, so that the agent expects that it will be pulled towards target states with a velocity proportional to the error.

Importantly, although the generalized belief allows encoding some information about dynamics of the true generative process, in this case the agent does not have any prior over it. For example, the agent does not know the trajectory of a moving target in advance (whose prior, in a more realistic scenario, would be certainly present and acquired through learning from past experiences), and will update the belief only relying on the incoming sensory information. Nevertheless, the agent maintains (false) expectations about target dynamics, and it is indeed the discrepancy between the evolution of the (real) generative process and that of the (internal and biased) generative model that makes the agent able to implement intention-driven behavior.

The prediction errors of the dynamics functions can be grouped in a single matrix:

$$\boldsymbol{E}^{(f)} = \boldsymbol{\mu}' - k\boldsymbol{E}^{(i)} \tag{27}$$

Thus, the derivative of free energy with respect to the belief is given by (also see Equation 14):

$$\dot{\tilde{\boldsymbol{\mu}}} = D\tilde{\boldsymbol{\mu}} - \partial_{\tilde{\mu}} F = \begin{bmatrix} \boldsymbol{\mu}' + \boldsymbol{G}^T(\boldsymbol{\pi} \odot \boldsymbol{e}_s) + (\boldsymbol{F} \odot \boldsymbol{E}^{(f)})\boldsymbol{\gamma}^T \\ \boldsymbol{\mu}'' - \boldsymbol{E}^{(f)}\boldsymbol{\gamma}^T \end{bmatrix} \tag{28}$$

Here $\odot$ is the element-wise product, $\boldsymbol{G}$ and $\boldsymbol{F}$ enclose the gradients of all sensory generative models and dynamics functions, while $\boldsymbol{\pi}$ and $\boldsymbol{\gamma}$ comprise all sensory and intention precisions:

$$\boldsymbol{G} = \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{\mu}} = \begin{bmatrix} \partial \boldsymbol{g}_0 \\ \vdots \\ \partial \boldsymbol{g}_L \end{bmatrix} \qquad \boldsymbol{\pi} = \begin{bmatrix} \pi_0 \\ \vdots \\ \pi_L \end{bmatrix} \qquad \boldsymbol{F} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\mu}} = \begin{bmatrix} \partial \boldsymbol{f}^{(0)} & \dots & \partial \boldsymbol{f}^{(K)} \end{bmatrix} \qquad \boldsymbol{\gamma} = \begin{bmatrix} \gamma^{(0)} & \dots & \gamma^{(K)} \end{bmatrix} \tag{29}$$

In the following we will neglect the backward error in the 0th-order of Equation (28) since it has a much smaller impact on the overall dynamics, and treat as the actual attractor force the forward error at the 1st-order:

$$\dot{\tilde{\boldsymbol{\mu}}} \approx \begin{bmatrix} \boldsymbol{\mu}' + \boldsymbol{G}^T(\boldsymbol{\pi} \odot \boldsymbol{e}_s) \\ \boldsymbol{\mu}'' - \boldsymbol{E}^{(f)}\boldsymbol{\gamma}^T \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}' + \boldsymbol{\epsilon}_s \\ \boldsymbol{\mu}'' - \boldsymbol{\epsilon}^{(i)} \end{bmatrix} \tag{30}$$

where $\boldsymbol{\epsilon}_s$ and $\boldsymbol{\epsilon}^{(i)}$ respectively stand for the contributions of precision-weighted sensory and intention prediction errors. We can interpret $\gamma^{(i)}$ as a quantity that determines the attractor gain of intention i, so that intentions with greater strength have a greater impact on the update direction; these gains could be modulated by projections from higher levels or frontal areas applying cognitive control, such as PFC or PMd. In turn, π_j corresponds to the confidence that the agent has about each sensory modality j, so that the agent relies more on sensors with higher strength.

Similarly, we can specify control signals by minimizing free energy with respect to actions, expressing the mapping from sensations to actions by:

$$\begin{aligned} \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{a}} &= \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{a}} \cdot \boldsymbol{G} \\ \dot{\boldsymbol{a}} &= -\partial_a F = -\partial_a \boldsymbol{\mu}^T \boldsymbol{\epsilon}_s \end{aligned} \tag{31}$$

where $\partial_a \boldsymbol{\mu}$ is an inverse model from belief to actions. If motor actions are specified as joint velocities, we can decompose and approximate the inverse model as follows:

$$\partial_a \boldsymbol{\mu} = \frac{\partial \boldsymbol{q}}{\partial \boldsymbol{a}} \cdot \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{q}} = \frac{\partial \boldsymbol{g}_p}{\partial \boldsymbol{a}} \cdot \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{g}_p} = \Delta_t \boldsymbol{G}_p^{-1} \tag{32}$$

where the subscript p indicates the proprioceptive contribution and we approximated $\partial_a \boldsymbol{g}_p$ by a time constant Δ_t (***Oliver et al., 2019***).

If we assume that the belief over hidden states is encoded in joint angle coordinates, the computation of the inverse model may be as simple as finding the pseudoinverse of a matrix; however, if the belief is specified in a more generic frame of reference and the proprioceptive generative model is a nonlinear function, it could be much harder to compute the corresponding gradient, introducing other control problems such as temporal delays on sensory signals (***Friston, 2011***).

Algorithm 1 Goal-directed Active Inference

Input: $\boldsymbol{i}, \boldsymbol{g}, \partial\boldsymbol{\mu}, \boldsymbol{\gamma}, \boldsymbol{\pi}, k, \Delta_t$

$\boldsymbol{\mu}, \boldsymbol{\mu}', \boldsymbol{\mu}'', \leftarrow InitializeBelief()$
while $t < T$ **do**
  $\boldsymbol{H} \leftarrow \boldsymbol{i}(\boldsymbol{\mu})$ ▷ Goal-directed intentions
  $\boldsymbol{E}^{(f)} \leftarrow \boldsymbol{\mu}' - k(\boldsymbol{H} - \boldsymbol{\mu})$
  $\boldsymbol{\epsilon}^{(i)} \leftarrow \boldsymbol{E}^{(f)}\boldsymbol{\gamma}^T$ ▷ Precision-weighted intention errors
  $\boldsymbol{p} \leftarrow \boldsymbol{g}(\boldsymbol{\mu})$ ▷ Sensory predictions
  $\boldsymbol{e}_s \leftarrow \boldsymbol{s} - \boldsymbol{p}$
  $\boldsymbol{\epsilon}_s \leftarrow \boldsymbol{G}^T(\boldsymbol{\pi} \odot \boldsymbol{e}_s)$ ▷ Precision-weighted sensory errors
  $\dot{\boldsymbol{\mu}} \leftarrow \boldsymbol{\mu}' + \boldsymbol{\epsilon}_s$ ▷ Belief update
  $\dot{\boldsymbol{\mu}}' \leftarrow \boldsymbol{\mu}'' - \boldsymbol{\epsilon}^{(i)}$
  $\dot{\boldsymbol{a}} \leftarrow -\partial\boldsymbol{\mu} \cdot \boldsymbol{\epsilon}_s$ ▷ Action update
  $\tilde{\boldsymbol{\mu}} \leftarrow \tilde{\boldsymbol{\mu}} + \Delta_t \dot{\tilde{\boldsymbol{\mu}}}$ ▷ Free energy minimization
  $\boldsymbol{a} \leftarrow \boldsymbol{a} + \Delta_t \dot{\boldsymbol{a}}$
end while

Alternatively, we can consider motor control driven only by proprioceptive predictions and sensations, so that the control signal is already in the correct domain and may be achieved through simple reflex arc pathways (***Adams et al., 2013***; ***Versteeg et al., 2021***). In this case, all that is needed is a mapping from proprioceptive predictions to actions:

$$\dot{\boldsymbol{a}} = -\partial_a F_p = -\Delta_t \cdot \pi_p \boldsymbol{e}_p \tag{33}$$

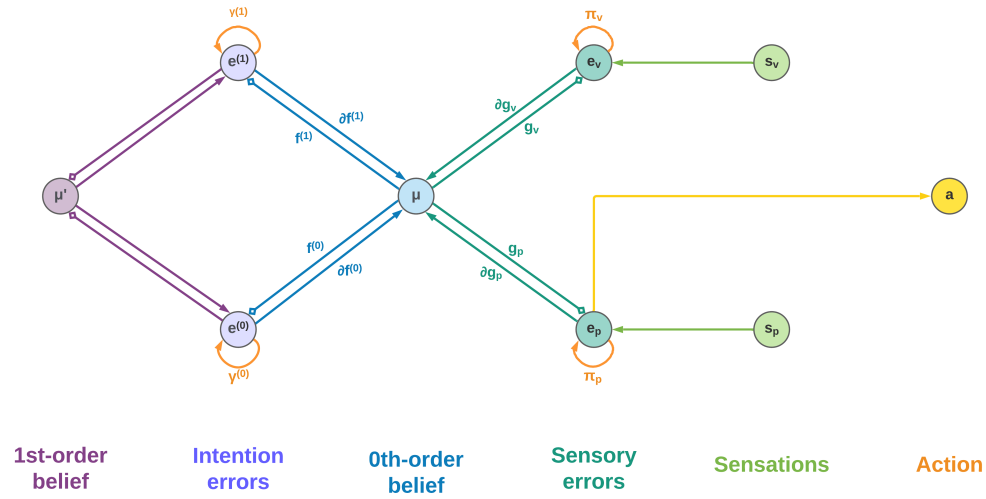


Figure 2. Neuronal representation of intention-mediated free energy minimization implementing Active Inference. Small squares stand for inhibitory connections.

Algorithm 1 outlines a schematic description of the flow of dynamic computations.

Neural Implementation

Figure 2 shows a schematic neuronal implementation of intention-mediated free energy minimization, which further extends earlier perceptual inference schemes (***Bogacz, 2017***) to full-blown Active Inference. For simplicity, the intentions consist only of a single layer with a neuron for each function, but they may be easily implemented with deeper neural networks. Note that intentions $\boldsymbol{h}^{(i)}$ and sensory generative models $\boldsymbol{g}_j$ are all part of the same architecture, the only difference being the location in the cortical hierarchy.

Low-level prediction errors for each sensory modality are represented by neurons whose dynamics depends on both the observations and the predictions of the sensory generative models:

$$\dot{\boldsymbol{e}}_j = \boldsymbol{s}_j - \boldsymbol{g}_j(\boldsymbol{\mu}) - \frac{\boldsymbol{e}_j}{\boldsymbol{\pi}_j} \tag{34}$$

Upon convergence of the neural activity, that is, $\dot{\boldsymbol{e}}_j = 0$, we obtain the prediction error computation derived above. In turn, the internal activity of neurons corresponding to high-level prediction errors is obtained by subtracting the generated intention from the 1st-order belief:

$$\dot{\boldsymbol{e}}^{(i)} = \boldsymbol{\mu}' - \boldsymbol{f}^{(i)}(\boldsymbol{\mu}) - \frac{\boldsymbol{e}^{(i)}}{\boldsymbol{\gamma}^{(i)}} \tag{35}$$

Having computed the information coming from the top and bottom of the hierarchy, the belief is updated by integrating every signal:

$$\dot{\boldsymbol{\mu}} = \sum_j \partial \boldsymbol{g}_j \boldsymbol{e}_j + \sum_i \partial \boldsymbol{f}^{(i)} \boldsymbol{e}^{(i)} \tag{36}$$

which parallels the update formula derived above (Equation 30).

Correspondingly, the 1st-order component of the belief is updated as follows:

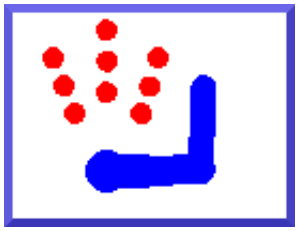
$$\dot{\boldsymbol{\mu}}' = -\sum_i \boldsymbol{e}^{(i)} \tag{37}$$

The belief is thus constantly pushed towards a direction that matches sensations on one side and intentions on the other. Since precisions are encoded in synaptic strengths, they can be dynamically optimized during inference in a direction that minimizes free energy (e.g., if a sensory modality is not useful to predict sensations, its weight will decrease). In this way, by dynamically

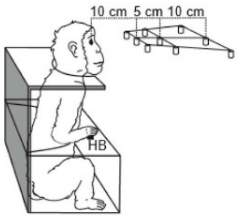
changing during the task, the intention weights can act as modulatory signals that select the best intention to realize at every moment.

Simulated Agent

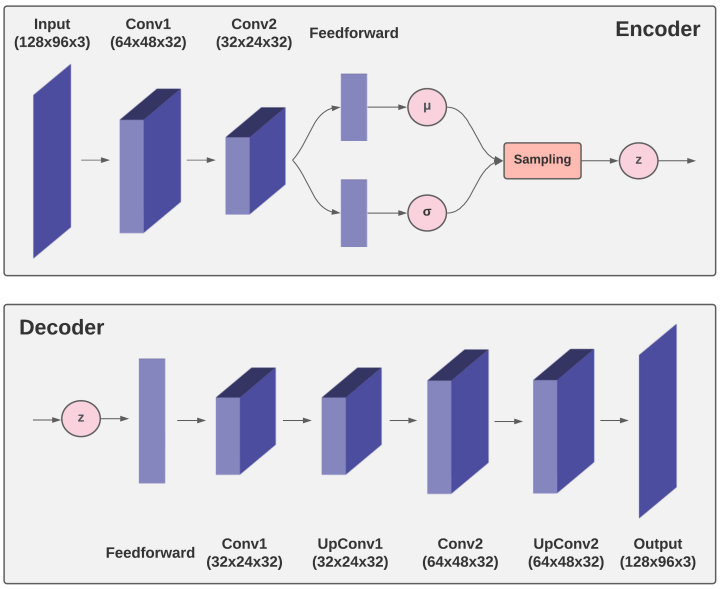
To demonstrate the feasibility of the approach, we simulated an agent consisting of an actuated upper limb with visual and proprioceptive sensors that allow it to perceive and reach static and moving targets within its reach. Figure 3a shows the size and position of the targets, as well as limb size and a sample posture. Since the focus here was on the theoretical approach, we simulated just a coarse 3-DoF limb model moving on a 2D plane. However, the approach easily generalizes to a more elaborated limb model and 3D movements.



(a) Simulation setup with all targets visible.



(b) Reference monkey experiment.



(c) Architecture of the variational autoencoder used in the simulation.

Figure 3. Outline. (a) Simulated experiment with an agent reaching for a red circle target as in the reference monkey experiment (***Breveglieri et al., 2014***) outlined in (b). (c) Variational autoencoder implementing (top) an inverse visual model, or encoder, and (bottom) a forward generative model, or decoder $\boldsymbol{g}_v$.

In the following, we first describe the agent and the specific implementation of the theoretical approach. Then, we provide an assessment of the agent's perceptual and motor control capabilities in static and dynamic conditions. The static condition simulated a typical monkey reaching task of peripersonal targets as in Figure 3b (***Breveglieri et al., 2014***). In turn, the dynamic condition involved a moving target that the agent had to track continuously.

Body

The body consists of a simulated monkey upper limb that includes a moving torso attached to an anchored neck, an upper arm, and a lower arm, as shown in Figure 3a. The three moving

segments are schematized as rectangles, each with unit mass, and the joints (shoulder, elbow) and the tips (neck, hand) as circles. The proportions of the limb segment and the operating range of the joint angles were derived from monkey data *Macaca mulatta* (***Kikuchi and Hamada, 2009***). The parameters of the limb model, including the rotation limits, are described in *Supplementary Material*, Table 1. The state of the limb and its dynamics are described by the joint angles $\boldsymbol{\theta}$ and their first moment $\dot{\boldsymbol{\theta}}$. We assume noisy velocity-level motor control, whereby the motor efferents $\boldsymbol{a}$ noisily controls the first moment of joint angles with zero-centered Gaussian noise:

$$\dot{\boldsymbol{\theta}} = \boldsymbol{a} + \boldsymbol{w}_a \tag{38}$$

Sensors

The agent receives information about its proprioceptive state and visual context. Simplified peripersonal visual input $\boldsymbol{s}_v$ was provided by a virtual camera that included three 2D color planes, each 128x96 pixels in size. The location and orientation of the camera were fixed so that the input provided full vision of the entire limb in any possible limb state within its operating range and of peripersonal targets. The limb could occlude the target in some configurations.

As in the simulated limb, the motor control system also receives proprioceptive feedback through sensors $\boldsymbol{s}_p$, providing noisy information on the true state of the limb (***Tuthill and Azim, 2018***; ***Versteeg et al., 2021***). We made a further simplification that $\boldsymbol{s}_p$ provides a noisy reading of the state of all joints only in terms of joint angles, ignoring other proprioceptive signals such as force and stretch (***Srinivasan et al., 2021***), which the Active Inference framework can natively incorporate.

Belief

We assume internal representations - belief $\boldsymbol{\mu}$ - approximating the latent causes $\boldsymbol{z}$ of the environment, including here three components: (i) beliefs $\boldsymbol{\mu}_a$ over arm joint angles, or posture (ii) beliefs $\boldsymbol{\mu}_t$ over the target location represented again in joint angle space, i.e., the posture corresponding to the arm touching the target, and (iii) beliefs $\boldsymbol{\mu}_h$ over a memorized home button (HB) position; thus, $\boldsymbol{\mu} = [\boldsymbol{\mu}_a, \boldsymbol{\mu}_t, \boldsymbol{\mu}_h]$. Note that the last two components can be interpreted as *affordances*, allowing the agent to implement interactions in terms of bodily configurations (***Pezzulo and Cisek, 2016***).

Intentions and Intention Functions

We recall that intentions encode desired future states of belief, produced from the current belief through intention functions. Stepping on the proposed formalization (Equation 20), we can define two specific intention functions (Figure 4) as follows:

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{i}^{(t)}(\boldsymbol{\mu}) & \boldsymbol{i}^{(h)}(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}^{(t)}\boldsymbol{\mu} & \boldsymbol{I}^{(h)}\boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_t & \boldsymbol{\mu}_h \\ \boldsymbol{\mu}_t & \boldsymbol{\mu}_t \\ \boldsymbol{\mu}_h & \boldsymbol{\mu}_h \end{bmatrix} \tag{39}$$

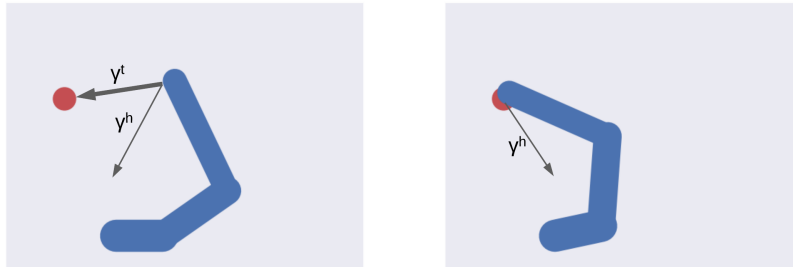
Here $\boldsymbol{h}^{(t)} = \boldsymbol{i}^{(t)}(\boldsymbol{\mu})$ defines the agent's expectation that the belief over arm joint angles is equal to the joint angle configuration corresponding to the target to be reached, and it is implemented as a simple mapping $\boldsymbol{I}^{(t)}$ that makes the first belief component equal to the second one. In turn, the intention $\boldsymbol{h}^{(h)} = \boldsymbol{i}^{(h)}(\boldsymbol{\mu})$ encodes the future belief of the agent that the arm will be in the HB position. The two intention mappings are defined by:

$$\boldsymbol{I}^{(t)} = \left[\begin{array}{c|c|c} \boldsymbol{0} & \boldsymbol{I}_3 & \boldsymbol{0} \\ \hline \boldsymbol{0} & \boldsymbol{I}_3 & \boldsymbol{0} \\ \hline \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_3 \end{array}\right] \qquad \boldsymbol{I}^{(h)} = \left[\begin{array}{c|c|c} \boldsymbol{0} & \boldsymbol{I}_3 & \boldsymbol{0} \\ \hline \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_3 \\ \hline \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_3 \end{array}\right] \tag{40}$$

where $\boldsymbol{0}$ and $\boldsymbol{I}_3$ are 3x3 zero and identity matrices.

The intention prediction errors for these two specific intentions are then:

$$\boldsymbol{E}^{(i)} = \begin{bmatrix} \boldsymbol{e}^{(t)} & \boldsymbol{e}^{(h)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}^{(t)} - \boldsymbol{\mu} & \boldsymbol{h}^{(h)} - \boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_t - \boldsymbol{\mu}_a & \boldsymbol{\mu}_h - \boldsymbol{\mu}_a \\ \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \tag{41}$$



(a) Reaching action. **(b)** Return action.

Figure 4. Multiple intentions (arrows) and relative strength (arrow width) with different attractors controlled by intention precisions. (a) Stronger attraction of the red circle implements the reach action. (b) Stronger attraction of the invisible but previously memorized HB implements the return-to-home action.

so to provide to the agent an update direction towards both the target joint angles and HB position. As there is no intention to move the target or the HB, the second and third components of the prediction errors will be zero. Importantly, this attractor is dynamically generated at each time step, so that the agent can also follow moving targets and interact with a constantly changing environment.

Sensory Predictions

The sensory generative density has two components, one for each sensory modality: a simplified proprioceptive model $\boldsymbol{g}_p(\boldsymbol{\mu})$ and a full-blown visual perceptual model $\boldsymbol{g}_v(\boldsymbol{\mu})$:

$$\boldsymbol{g}(\boldsymbol{\mu}) = \begin{bmatrix} \boldsymbol{g}_p(\boldsymbol{\mu}) \\ \boldsymbol{g}_v(\boldsymbol{\mu}) \end{bmatrix} \tag{42}$$

Since here the belief is already in the joint angle domain, we implemented a simple proprioceptive generative model $\boldsymbol{g}_p(\boldsymbol{\mu}) = \boldsymbol{G}_p\boldsymbol{\mu} = \boldsymbol{\mu}_a$, where $\boldsymbol{G}_p$ is a mapping that only extracts the belief over arm joint angles from the overall belief space:

$$\boldsymbol{G}_p = \left[\begin{array}{c|c|c} \boldsymbol{I}_3 & \boldsymbol{0} & \boldsymbol{0} \end{array} \right] \tag{43}$$

Note that $\boldsymbol{g}_p(\boldsymbol{\mu})$ could be easily extended to a more complex proprioceptive mapping if the body and/or joint sensors have a more complex structure and the belief has a richer and more abstract representation.

In turn, the visual generative model $\boldsymbol{g}_v$ is the decoder component of a VAE (see Figure 3c) that produces a visual representation of both target and body configuration given the overall belief in joint angles (example in Figure 15a). The decoder consists of one feedforward layer, two transposed convolutional layers, and two other standard convolutional layers needed to smooth the output. All VAE parameters are described in *Supplementary Material*, Table 2.

The gradient of the full generative model $\boldsymbol{G}$ is the following:

$$\boldsymbol{G} = \begin{bmatrix} \partial \boldsymbol{g}_p \\ \partial \boldsymbol{g}_v \end{bmatrix} \tag{44}$$

Here, the proprioceptive gradient simplifies to the proprioceptive mapping $\boldsymbol{G}_p$ itself, while the visual component is the gradient of the decoder computed by backpropagation.

Thus, the prediction $\boldsymbol{p}$ and the prediction error $\boldsymbol{e}_s$ take the form:

$$\boldsymbol{p} = \begin{bmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{g}_v(\boldsymbol{\mu}) \end{bmatrix} \qquad \boldsymbol{e}_s = \begin{bmatrix} \boldsymbol{s}_p - \boldsymbol{\mu}_a \\ \boldsymbol{s}_v - \boldsymbol{g}_v(\boldsymbol{\mu}) \end{bmatrix} \tag{45}$$

Note that defining the sensory predictions on both proprioceptive and visual sensory domains allows the agent to perform efficient goal-directed behavior also in conditions of visual uncertainty,

e.g., due to low visibility. Indeed, since the belief is maintained over time until update, the agent maintains the last known target position and can thus accomplish reaching tasks also in case of temporarily occluded targets.

Precisions

Free energy minimization and predictive coding in general heavily depends on precisions. We adopted the idea that the slow-varying precisions are encoded as synaptic strengths (***Bogacz, 2017***); Figure 2. Alternative ideas see them as gains of superficial pyramidal neurons (***Bastos et al., 2012***).

To investigate their role, we parameterized the relative precision of each intention and sensory domain with parameters α and β as follows:

$$\boldsymbol{\pi} = \begin{bmatrix} 1-\alpha \\ \alpha \end{bmatrix} \qquad \boldsymbol{\gamma} = \begin{bmatrix} 1-\beta & \beta \end{bmatrix} \tag{46}$$

The parameter α controls the relative strength of the error update due to proprioception and vision, while the parameter β controls the relative attraction over the belief by each intention. With these parameters, the sensory and intention prediction errors are unpacked as follows:

$$\boldsymbol{\epsilon}_s = (1-\alpha) \cdot \boldsymbol{G}_p^T \boldsymbol{e}_p + \alpha \cdot \partial \boldsymbol{g}_v^T \boldsymbol{e}_v \tag{47}$$

$$\boldsymbol{\epsilon}^{(i)} = \boldsymbol{\mu'} - k[(1-\beta) \cdot \boldsymbol{e}^{(t)} + \beta \cdot \boldsymbol{e}^{(h)}] \tag{48}$$

Equation (47) shows the balance between visual and proprioceptive information. For example, if $\alpha = 0$, the agent will only use proprioceptive feedback, while if $\alpha = 1$, the belief will be updated only relying on visual feedback. Note that these are extreme conditions (e.g., the former may correspond to null visibility), and typical sensory systems provide balanced feedback.

In turn, Equation (48) spells out the control of belief attraction. When $\beta = 0$, the agent will follow the first intention, while it will follow the second intention when $\beta = 1$ (Figure 4). Intermediate attractor strengths (β values) will lead to an attractor conflict, thus we assume that biased competitive inhibition implements the control of intention selection. Finally, the parameter k controls the overall attractor magnitude (see also Equation 26).

We can also use the precision parameter α to manipulate the strength of the derivative of free energy with respect to action (i.e., motor control) as follows:

$$\dot{\boldsymbol{a}} = -\Delta_t(1-\alpha) \cdot \boldsymbol{e}_p \tag{49}$$

Note that by increasing α (i.e., more reliability on vision), the magnitude of belief update remains constant, while action updates decrease because the agent becomes less confident about its proprioceptive information. Also, one could differentially investigate the effect of precision strength on belief and action by directly manipulating specific precisions; e.g., visual precisions $\boldsymbol{\pi}_v$ may include components following the belief structure (*Supplementary Material*, Table 3):

$$\boldsymbol{\pi}_v = [\alpha, \pi_{v_t}, \pi_{v_h}] \tag{50}$$

For example, when $\alpha = 0$ and $\pi_{v_t} > 0$, belief on the target is updated using visual input while the arm moves only using proprioception, a scenario that emulates movement in darkness with a lit target.

Computational Analysis

In the following, we describe an assessment of the capacity of the intention-driven Active Inference agent to perceive and perform goal-directed actions in reaching tasks with static and dynamic visual targets. The main testbed task was delayed reaching, but we also assessed other conditions.

Sensorimotor control that implements goal-directed actions was investigated in various sensory feedback conditions, including pure proprioceptive, also known as open loop, and mixed visual and proprioceptive, or closed loop (***Todorov, 2004***), in which the VAE decoder provided support for dynamic estimation of belief over visual targets and bodily states and generated visual

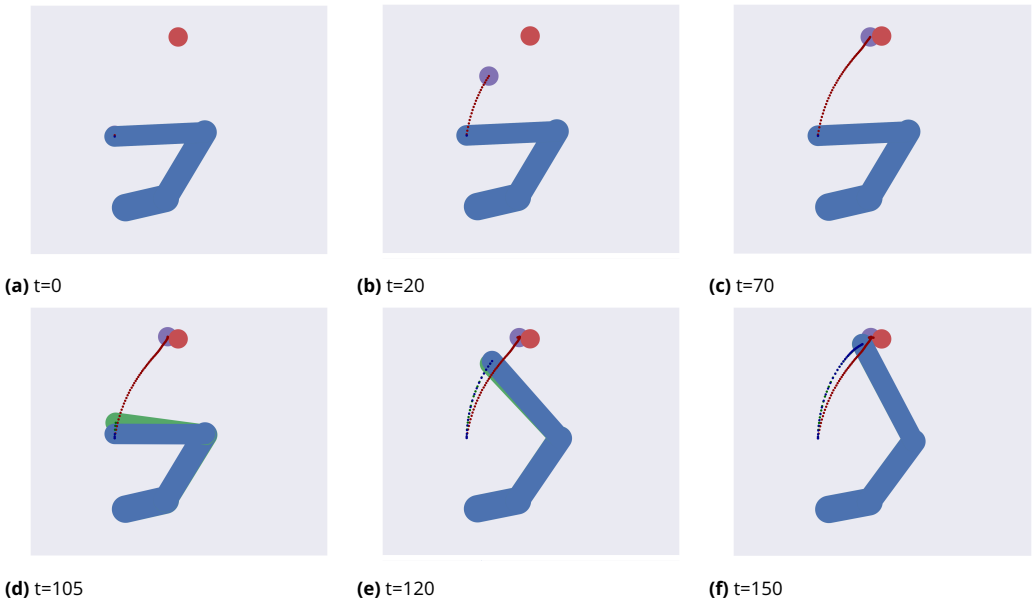


Figure 5. Delayed reaching. At trial onset (t=0), visual target (red circle) appears, arm (in blue) is located on the HB position, and belief over arm (in green) is set at the true arm state. During the delay period the iterative perceptual inference process gradually drives target belief (purple circle) towards the target (t=20 and t=70). During this phase the intention gain k is set to 0, so movement is inhibited and arm belief does not change given the unchanged proprioceptive evidence (t=70). Upon movement deliberation, the arm freely follows arm belief (t=105 and t=120) until they both arrive at destination (t=150).

sensations. In an additional *baseline* (BL) condition, the target was estimated by the decoder, but movement was performed without visual feedback or proprioceptive noise, in order to allow comparison with the typical approach in previous dynamic Active Inference studies, e.g., (***Pio-Lopez et al., 2016***). We also investigated the effects of sensory and intention precisions, motor control type, and movement onset policy. Finally, we analyzed the visual model and the nature of its gradients to provide critical information about the causes of the observed motor behavior.

Action performance was assessed with the help of several measures: (i) *reach accuracy*: success in approaching the target within 10 pixels of its center, that is, the hand touching the target; (ii) *reach error*: L^2 hand-target distance at the end of the trial; (iii) *reach stability*: standard deviation of L^2 hand-target distance during the period from target reach to the end of the trial, in successful trials; (iv) *reach time*: number of time steps needed to reach the target in successful trials.

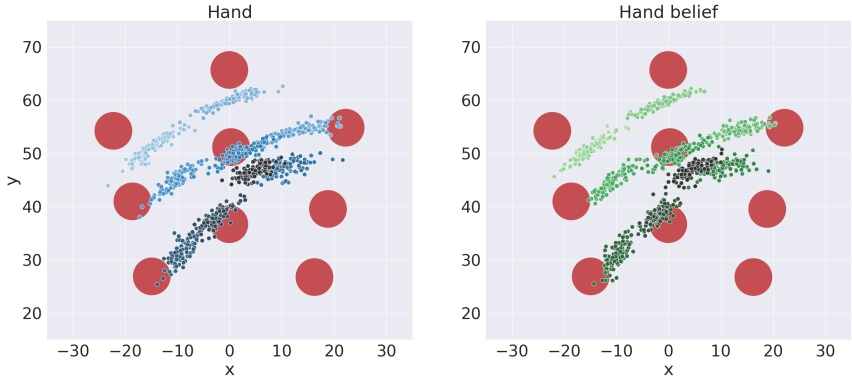
We also assessed target perception through analog measures based on the L^2 distance between the target location and its estimate transformed from joint angle coordinates into visual position by applying the geometric (forward) model. Specifically, we defined the following measures: (v) *perception accuracy*: success in estimating the true target location within 10 pixels; (vi) *perception error*: L^2 distance between the true and estimated target position at the end of the trial; (vii) *perception stability*: standard deviation of the L^2 distance between the target position and its estimation during the period starting from successful estimation until trial end; (viii) *perception time*: number of time steps needed to successfully estimate the target position.

Delayed Reaching Task

The primary testbed task is a simplified version of a delayed reaching monkey task in which a static target must be reached with a movement that can only start after a delay period (***Breveglieri et al., 2014***). Delayed actions are used to separately investigate neural processes related to action preparation (e.g., perception and planning) and execution in goal-directed behavior. Here, the task allows for the separate investigation of the two main computational components of free energy minimization, namely, perceptual and active inference, which otherwise work in parallel. Delayed

reaching could be implemented using various approaches: the update of the posture component of belief dynamics could be blocked by setting during inference the corresponding intention gain k to 0 (implemented here): in this way there are no active intentions and the belief only follows sensory information; alternatively, action execution could be temporarily suspended by setting the parameter α to 1, so that the agent still produce proprioceptive predictions but does not trust the corresponding prediction error: in this scenario, the belief dynamics include a small component directed towards the intention, but the discrepancy produced is not minimized through movement.

(a) Final positions with visual feedback.

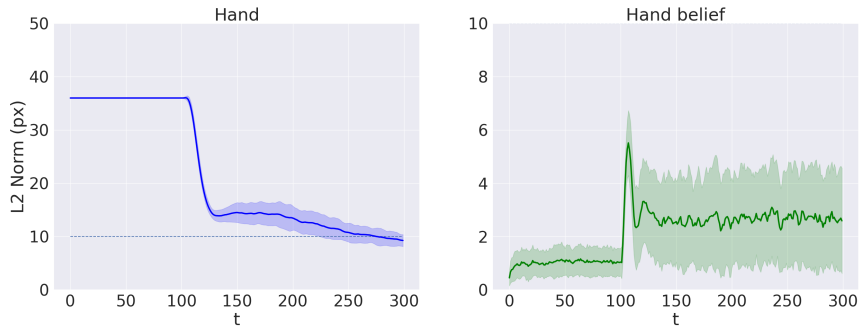


(b) Final positions without visual feedback.

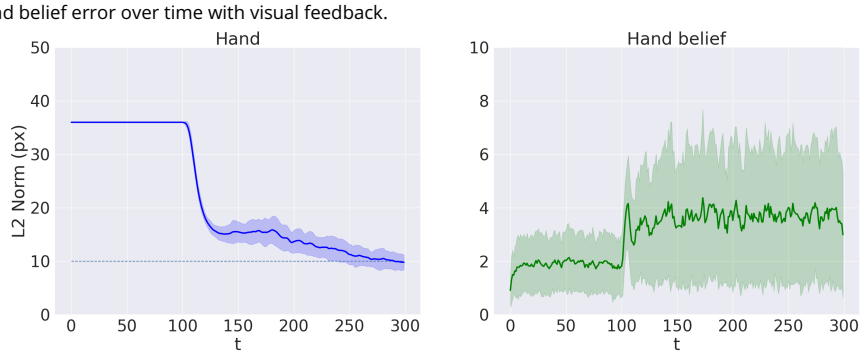
Figure 6. Spatial distribution of hand positions (left) and corresponding belief (right) per target at the end of reach movements, with (a) and without (b) visual feedback. Each point represents a trial (100 trials per target).

Reach task trials start with the hand placed on an HB located in front of the body center (i.e., the "neck") and belief is initialized with this configuration. Then one of 9 possible targets similar to those of the reference experiment (Figure 3b) is lit red. Follows a delay period of 100 time steps during which the agent is only allowed to perceive the visible target and the limb, i.e., the inference process can only change belief. After that, the limb is allowed to move, that is, the joint angles are updated according to (38). As in the reference task, upon reaching the target, the agent maintains the hand on the target for a sufficiently long period, for a total of 300 time steps per trial. After that, the agent reaches back the HB (not analyzed here). The simulation included 100 repetitions per target, that is, 900 trials in total.

As described before, we assume that the HB position is encoded and memorized as a belief during previous experience, and therefore used to compute the (competing) reaching intention. However, this creates a conflict among intentions aiming to fulfil opposing goals, while the agent can physically realize only one of them at a time (Figure 4). To fulfil one competing intention at a time (e.g., reach first a target and then the HB), we could assume that intention precisions are dynamically modulated through high-level (e.g., prefrontal) bias and mutual inhibition. We imple-



(a) Reach and belief error over time with visual feedback.



(b) Reach and belief error over time without visual feedback.

Figure 7. Performance on the delayed reaching task with (a) and without (b) visual feedback (100 trials per condition; bands represent C.I.). The reach criterion of hand-target distance less than 10 pixels is visualized as a dotted line. Reaching with visual feedback resulted in a more stable hand belief. L^2 norm for hand belief is computed by the difference between real and estimated hand position.

mented this with the parameter β (Equation 48), which was set to the more realistic values of 0.1 to activate the reaching intention and 0.9 to implement the intention to return to the HB.

Figure 5 illustrates the overall dynamic Active Inference process of perception and motor control that implements the delayed reaching task. Following target presentation (panel a), during the delay period (panels b-c), posture (blue arm) does not change since joint angles only follow the belief over arm (in green), which is now kept fixed, while belief over target (purple circle) is attracted by the sensory evidence (red circle) and gradually shifts towards it. As explained above, in our novel intention-mediated model, actions are driven by an attractor that combines the joint effect of all intention functions selectively activated by the β parameter. Thus, when movement is allowed (panels d-f), by setting $\beta = 0.1$, the combined attractor produces a force that moves the arm belief toward the target, which in turn generates proprioceptive predictions and therefore motor commands that pull the arm toward the target.

Figures 6a and 6b show spatial statistics of the final hand location and hand belief for each target, separately for reaching with proprioception only (open loop) or proprioceptive and visual (closed loop) sensory feedbacks. Descriptive statistics revealed an important benefit of visual feedback: reach accuracy was higher (with: 88.28%; without: 83.72%) and both reach stability and arm belief error were considerably better with visual feedback as well (stability: 1.35; error: 1.98px) compared to the condition with only proprioception (stability: 1.78; error: 2.87px). Therefore, in our model, closed loop control of reaching resulted in a more precise behavior (Figure 7), which parallels behavioral data (***Keele and Posner, 1968***).

Precision Balance

The effect of sensory feedback led to a further more systematic assessment of the effects of sensory and intention precisions, denoted by α, π_{v_t} and k (see Equations 46 and 50). The assessment

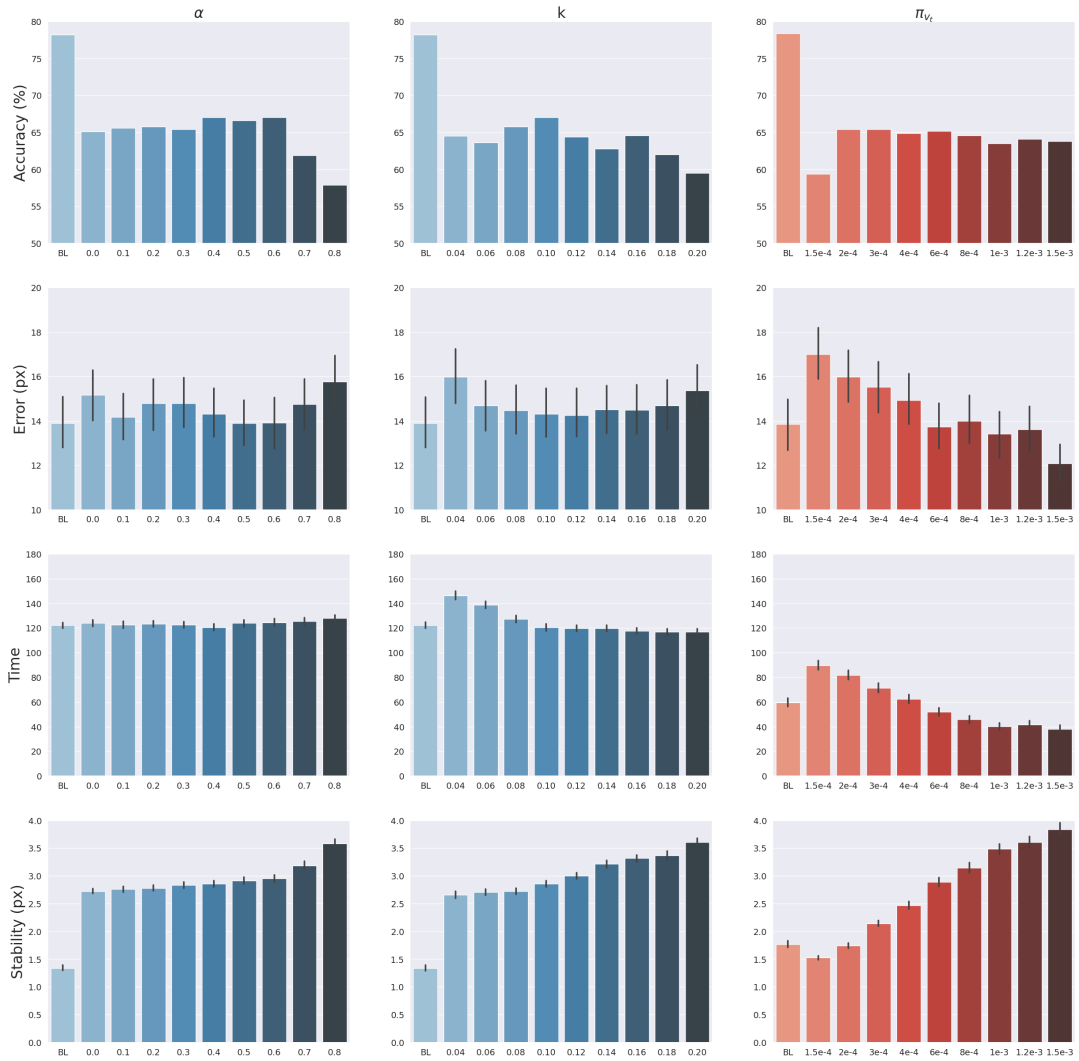


Figure 8. Effects of precision balance on reach accuracy (1st row), reach error (2nd row), i.e., L^2 hand-target distance, reach time (3rd row), and reach stability (4th row); for definitions, see text. Reach performance is shown as a function of arm sensory precision α (left) and attractor strength k (middle). In turn, perception performance is shown as a function of sensory precision over target π_{v_t} (right). Vertical bars represent C.I.

was carried out on a task that followed the structure of the main delayed reaching task. We varied the above precisions one at a time, using levels shown on the abscissas in Figure 8, while keeping the non-varied precisions at their default values. Note that the arm precision of $\alpha = 0$ corresponds to reaching without visual feedback, while the conditions $\alpha > 0$ correspond to reaching with different levels of arm visibility. We recall that the baseline condition (BL) performs reaching without visual feedback and proprioceptive noise, that is, $\alpha = 0$ and $w_p = 0$. To obtain a systematic evaluation, each precision condition was run on a rich set of 1000 randomly selected targets that covered the entire operational joint angle space. Finally, we only considered the target reaching intention, i.e., $\beta = 0$; everything else was the same as in the main task.

The results are shown in Figure 8. The panels in the left column show the effect of arm precision α compared to the BL agent with noiseless proprioception. Active Inference with only proprioception (i.e., $\alpha = 0$) has a lower reach accuracy and higher error, in corroboration with the observations in the basic delayed reaching task; the best performance is obtained with balanced proprioceptive and visual input, which is the typical condition of movement with visual feedback and generally allows optimal behavior (***Keele and Posner, 1968***). Indeed, in that case, the motor control circuitry

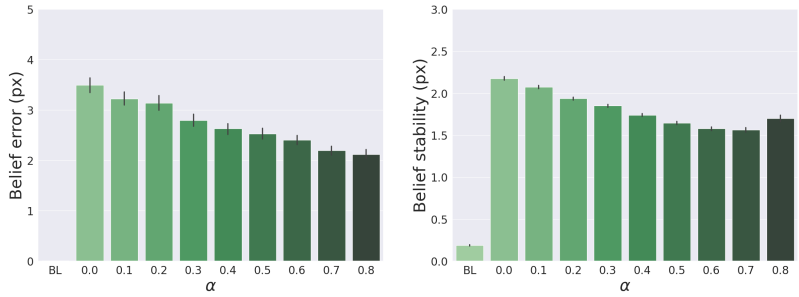


Figure 9. Belief error and stability (representing the difference between real and estimated hand position) for different values of α. Vertical bars represent C.I.

continuously integrates all available sensory sources to implement visually guided behavior (***Saunders and Knill, 2003***). However, accuracy and stability rapidly decrease for excessively high values of α, due to the discrepancy in update directions between belief, which makes use of all available sensory information, including increasingly more precise visual feedback, and action, which in this case relies on excessively noisy proprioception.

Furthermore, as in the main delayed reaching experiment, the effect of visual precision manipulation is also evident in the stability of arm belief, which gradually improves with the increase in visual input precision (Figure 9): In addition to the reliability of the visual input, this effect is also a consequence of the smaller action updates due to the reduced proprioceptive precision.

In turn, the panels in the middle column of Figure 8 reveal the effect of attractor gain k; to remind the reader, the greater the gain, the greater the contribution of intention prediction errors to the belief updates. The results show that as the intention gain k increases, reach accuracy generally improves and the number of time steps needed to reach the target decreases. However, beyond a certain level, the accuracy tends to decrease since the trajectory dynamics becomes unstable; thus, excessively strong action drag is counterproductive to the implementation of smooth movements.

Finally, the panels in the right column of Figure 8 show the effect of target precision π_{v_t}, which directly affects the quality of target perception. Note that contrary to the effect of arm visual precision, better performance is generally obtained in terms of accuracy, error, and perception time for higher values of π_{v_t}, which corresponds to a classical effect of contrast on perception, but also means here that arm and target belief updates follow different dynamics. Importantly, perception stability decreases as target precision levels increase, which deserves further exploration.

Motor Control

As explained earlier, motor control in Active Inference could be realized in two different ways: by making use of either all sensory information, thus, the total sensory prediction error including proprioception and exteroception (here, vision), or only the proprioceptive domain. The first method requires significantly more computations since the agent needs to know the inverse mapping from every sensory domain (generative model) in order to compute the control signals; however, considering the assumptions we made, this approach is more stable because it updates both belief (state estimation) and action with the same information. On the other hand, in the second proprioception-based control mechanism, the proprioceptive predictions are compared with the noisy proprioceptive observations, which produce potentially incorrect movements.

Figure 10 shows a comparison of the performance of the two motor control methods and the BL agent evaluated under the test condition used to investigate precision balance, including 1000 random targets. We here measured belief and reach stability and reach accuracy. The results reveal, first, that the decreased belief stability of the full model with respect to the BL agent (the right panel) does not affect hand stability appreciably (the middle panel), although the proprioceptive noise apparently contributes to decreased reach accuracy (see the left panel). More importantly, the results reveal that the proprioception model has considerably lower reach stability, which was

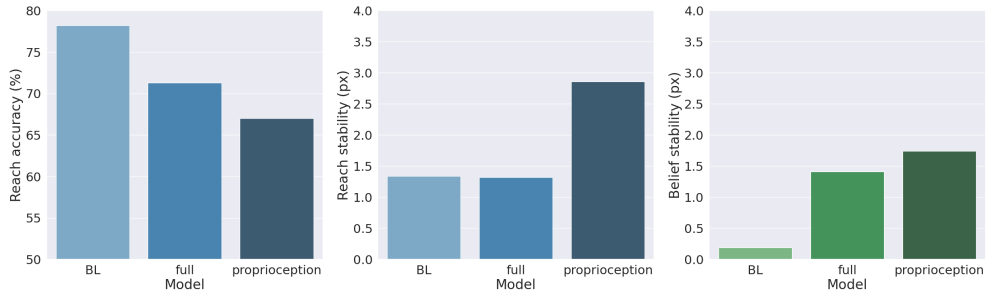


Figure 10. Motor control methods. Reach accuracy (left panel), reach stability (middle panel) and belief stability (right panel) for a BL agent and the two different implementations of motor control, based either on all sensory information (full control) or on proprioception only (proprioceptive control).

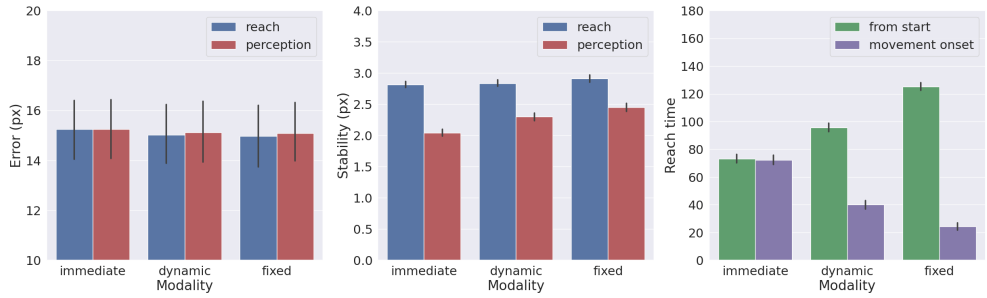


Figure 11. Effect of movement onset policy on error (left), stability (middle) and reach time (right) across several policies (immediate, dynamic, fixed delay). Vertical bars represent C.I.

caused by incorrect update directions of the motor control signals and which further contributes to decreasing reach accuracy with respect to the full model.

Movement Onset Policy

Furthermore, we investigated the effect of movement onset using several policies, which differ by the duration of the period of pure perception preceding full Active Inference, i.e., perception and action. Note that a relatively long pure perception period is inherent in the delayed reaching task that we used so far, which differs from a typical movement-onset regimen. One such policy that we investigate here is characteristic of actions performed under time pressure, in which movement starts along with perception, that is, action is *immediate*. Another policy that could be considered typical for acting under normal conditions has movement beginning with the satisfaction of a certain perception criterion. This policy *dynamically* deliberates the onset of movement; various perception criterion could be used: here, we used one according to which the norm of target belief $\dot{\boldsymbol{\mu}}_t$ remained below a given threshold (i.e., 0.01) for a certain period of time (i.e., 5 time steps). These parameters were arbitrarily chosen in consideration of exploratory delayed reaching simulations. Finally, we include the previously used delayed action policy in which movement onset is delayed by a *fixed* time period (here, 100 time steps, sufficient to obtain a precise target estimation). To obtain systematic observations, each policy was run on 1000 randomly selected reach targets also in these simulations. Measurements included reach and perception accuracy, motor control stability after reach, and target perception stability, as well as reach time since the beginning of the trial or movement onset.

Figure 11 shows the result of the assessment in the three different policies. Although the reach error is approximately the same in all tested conditions, the agent controlled by the immediate motion-start policy reached the target within the lowest total number of time steps: target perception and intention setting were dynamically computed along with movement onset. However, if we consider the total task time as a movement time, the number of time steps is the highest in this

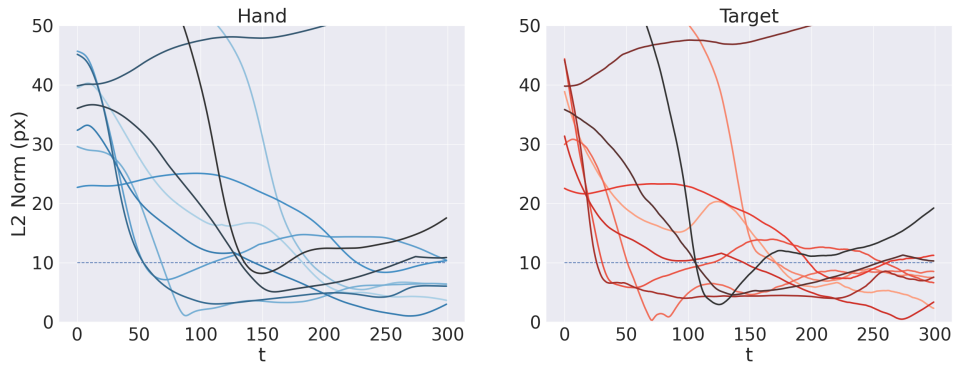


Figure 12. Dynamic targets: Reach (left) and perception (right) error over time for 10 random trials.

condition, since arm belief and the arm itself move along with the slow visual target estimation.

In turn, if the agent starts the movement when the uncertainty about the target position is already minimized (either in the *dynamic* or *fixed* condition), the movement time decreases, although the movement time added on top of the perception time (or delay period) results in slower actions relative to the immediate movement condition. Finally, we note that target (perception) stability somewhat decreases for dynamic and fixed policies; this somewhat unexpected result is encouraging for dynamic target tracking tasks in which immediate movement onset is mandatory.

Tracking Dynamic Targets

In a second testbed task, the agent was required to track a smooth moving target whose initial location was randomly chosen from the entire operational joint angle space. In each trial, the targets received an initial velocity of 0.1px per step in direction uniformly spanning the 0-360 degree range. When the target reached a border, its movement was reflected. As in the previous simulations, the belief was initialized at the HB configuration and the trial time limit was 300 time steps. However, in order for the agent to correctly follow the targets, belief and action were dynamically and continuously inferred in parallel, that is, without a perception period.

Figure 12 shows the reach trajectory in dynamic target tracking in 10 random trials, where each line is a trial. The left panel (a) shows the evolution over time of L^2 hand-target distance (reach error), while the right panel (b) represents the error between the estimated and the true position of the target (perception error). The results suggest that the agent is generally able to correctly estimate and dynamically keep updated the belief over both target and arm configuration for almost every trial, also in the case of moving targets. In some cases, though, mainly when the target is out of reach, it is temporarily or permanently "lost" in terms of its belief, which also has the consequences of losing the target in terms of reach. Further analysis with a more realistic bodily configuration and visual sensory system and comparison with actual kinematic data should provide further insights on the capabilities of Active Inference to perform dynamic reaching.

Free Energy Minimization

Figure 13 shows the trajectory of the free energy derivatives with respect to the arm component of generalized belief (upper two rows, corresponding to the sensory and intention prediction errors, respectively) and target belief (bottom row) in delayed reaching of a static target. The two columns correspondingly show the trends for the last two joints (the arm and forearm segments) that most strongly articulate the reach action.

Note that the gradients of the free energy with respect to the target components of the belief are rapidly minimized during the initial perceptual phase while the arm gradients remain still. Upon action execution, also the arm gradients rapidly change, computing new proprioceptive predictions that produce arm movements, and unexpectedly causing perceptual disturbance and increased free energy on target belief, which is rapidly minimized again.

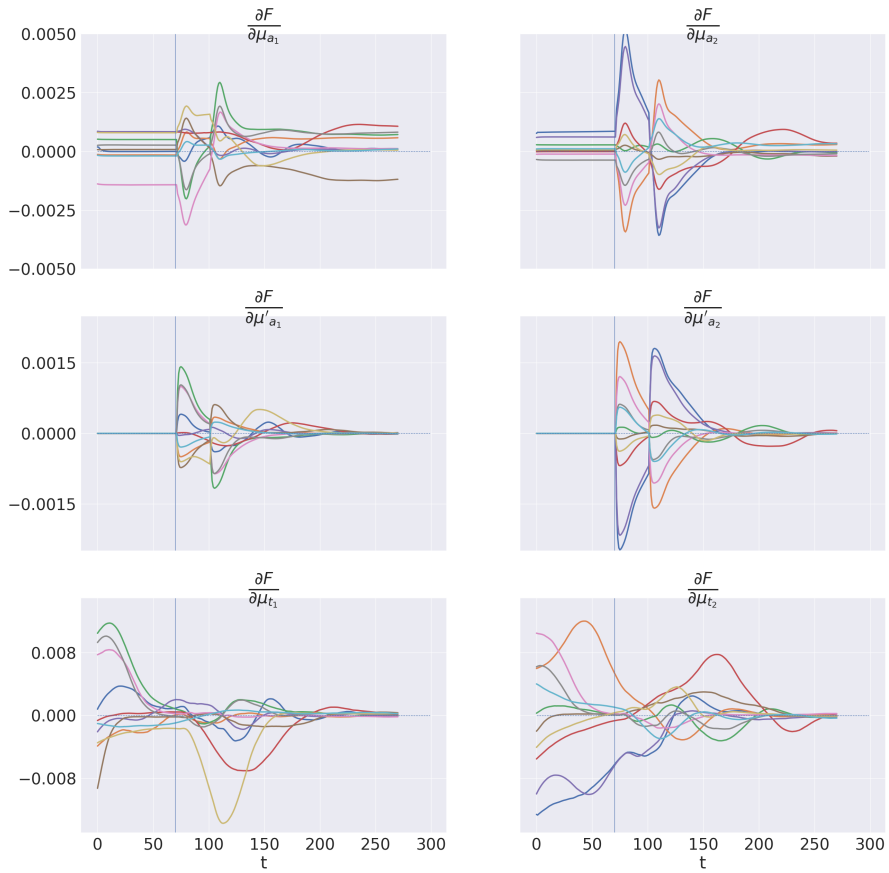


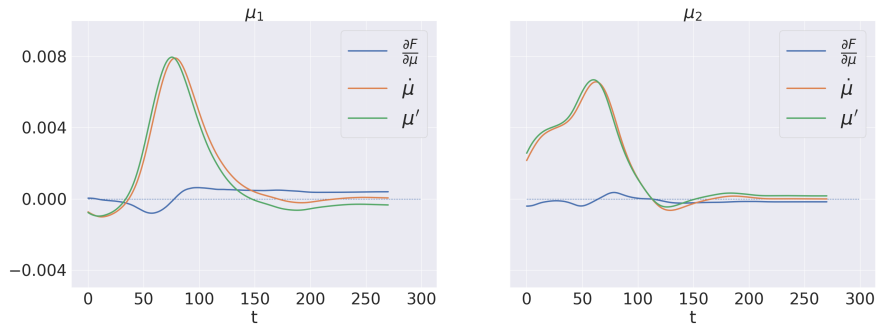
Figure 13. Free energy derivative with respect to the 0th- and 1st-order of generalized belief. Trials data is smoothed with a 30 time-step moving average. The left/right columns refer to arm/forearm.

Here we illustrate the dynamics of free energy minimization in reaching, which is at the heart of dynamic Active Inference. To that aim, we run new reaching trials with static and dynamic targets and recorded the derivatives of free energy with respect to generalized belief and action.

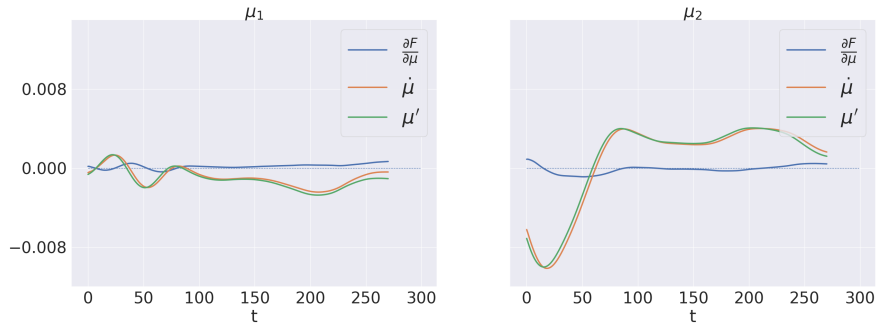
Figure 14 goes even deeper, showing a direct comparison between $\dot{\boldsymbol{\mu}}$, $\boldsymbol{\mu}'$, and $\dot{\boldsymbol{\mu}} - \boldsymbol{\mu}'$ on a sample static and dynamic target reaching trial. We recall that free energy minimization implies that the two frames of reference (the path of the mode $\dot{\boldsymbol{\mu}}$ and the mode of the path $\boldsymbol{\mu}'$) should overlap at some point in time, when the agent has inferred the correct trajectory (i.e., belief) of the generalized hidden states. This is crucial especially in dynamic reaching, in which the aim is to capture the instantaneous trajectory of every object in the scene through encoding in terms of generalized belief coordinates. The decreasing free energy gradients (blue lines) show that this aim is indeed successfully achieved in both static and dynamic tasks.

Visual Model Analysis

Here, we provide an assessment of the visual sensory model whose performance is critical for accurate visually guided motor control. To recall, the visual model is implemented with a VAE trained offline to reconstruct images of arm-target configurations such as the one in Figure 15a, left. A critical VAE parameter is the variance of the recognition (encoder) density Σ_ϕ (see Equati 17), therefore, we evaluated its effect on perception and action by training several VAEs with different variance levels. VAEs were trained for 100 epochs on a dataset comprising 20.000 randomly drawn body-target configurations that uniformly spanned the entire operational space. The target size varied, with a radius ranging from 5 to 12 pixels. VAE performance was assessed on other 10.000 randomly selected configurations that uniformly sampled the space, with target size of 5 pixels



(a) Static targets.



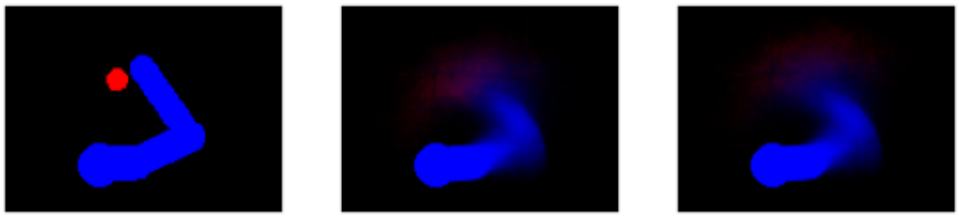
(b) Dynamic targets.

Figure 14. Comparison between frames of reference of belief for sample static (a) and dynamic (b) trials. The left/right columns refer to the arm/forearm segments.

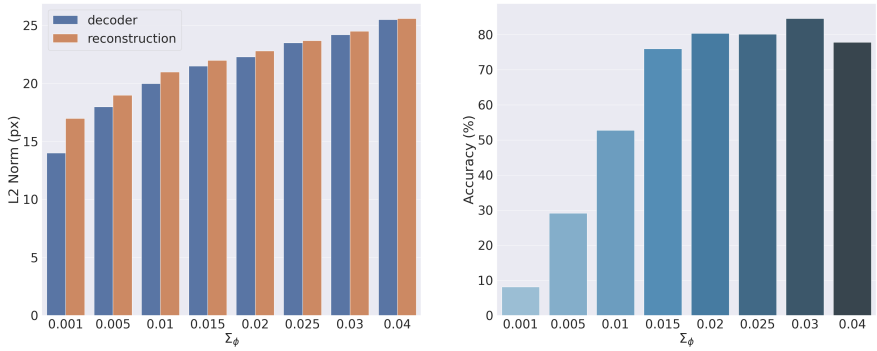
(the default condition for the Active Inference tests).

The most critical was the VAE capacity to generate adequate images of joint arm-target configurations, which we measured with the help of the L^2 norm in the visual domain between visual observations and VAE-generated images. To provide more insights on the two VAE processes, decoding and encoding, we proceeded as follows: first, *decoding* was assessed by generating images for given body-target states such as that in the middle of Figure 15a. The decoded images were compared with the ground truth images produced by applying the geometric model for the same state of the body target (Figure 15a, left). Second, full VAE performance was assessed by computing the average L^2 norm between observed images and their full VAE *reconstruction*, i.e., first encoding and then decoding (Figure 15a, right). Third, we assessed the effect of recognition density variance on active inference using as a measure the performance in the Baseline condition of the delayed target-reaching task.

Figure 15b, left, represents the results of the perceptual assessment tests, showing the L^2 norm between the original and generated images as a function of recognition density variance. As expected, lower variance generally resulted in lower error with respect to both pure decoding and full encoding-decoding (reconstruction). Surprisingly, the accuracy of active inference in the reaching task behaved differently: the best accuracy was obtained not for predictions with low variance, but for intermediate variance levels (Figure 15b, right). This could be explained with the fact that low-variance images imply highly nonlinear gradients that prevent correct gradient descent on free energy, while as the variance increases, the reconstructed image becomes somewhat blurred, which helps to obtain a smoother gradient that correctly drives free energy minimization and therefore improves movement accuracy. However, as the variance continues to increase, the reconstructed images become too blurry, degrading belief inference and motor control.



(a) Sample visual observation $\boldsymbol{s}_v$ (left); image decoded from joint angles (middle); reconstructed image (right).



(b) L^2-norm between observed and reconstructed image (left), and accuracy of active inference simulation (right).

Figure 15. Analysis of the perceptual model. (a) Sample image and its decoding from joint angles and through complete encoding-decoding process. (b) Quality of perception (left) and active inference (right) as a function of recognition density variance Σ_ϕ.

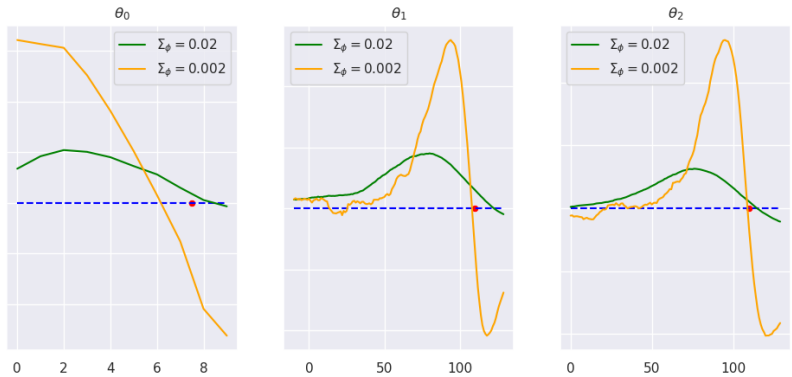
Visual Gradient Analysis

To further investigate the cause of the unexpected low variance issue, we analyzed the consistency of the visual gradient $\partial \boldsymbol{g}_v$ of the decoder for several encoder variance levels. To this aim, we computed the gradients for different reference (belief) states over the entire operational space.

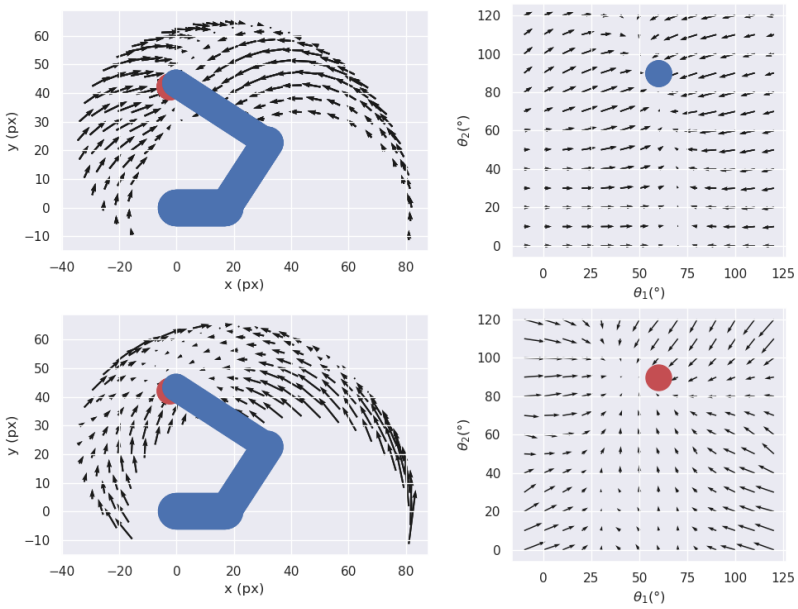
First, Figure 16a reveals that a decoder with medium variance levels (green line) causes smaller but smoother gradients, while too low variance (orange line) causes sharp peaks near the reference point and even incorrect gradient directions in some cases. It thus seems that too low encoder variance makes the decoder prone to overfitting, while higher variance levels help to extract a smoother relation between irregular multidimensional sensory domains and regular low-dimension causes.

Figure 16b further illustrates the arm and target gradients relative to a sample reference posture and target location (the result is similar for other configurations), in both Cartesian and polar (joint angle) coordinates; the polar plot shows the two joints most relevant to the reach action.

The plots reveal greater arm gradients (upper panels) in the vicinity of the reference location; in that subspace the decoder has less uncertainty about which direction to choose so to minimize the error. Notably, the set of gradients tends to compose not linear but curved directions, which is characteristic of biological motion. The polar plot provides critical insights into the cause of the circular pattern: the gradients are mostly parallel to the horizontal axis, which corresponds to a movement consisting essentially of pure shoulder rotation. Thus, the gradients provide a strong driving force on the shoulder almost throughout the operational space, while the area in which the elbow is controlled is limited to the vicinity of the reference point (e.g., target location). These gradients result in a two-phase reaching of static targets in which the agent first rotates the shoulder (implementing a horizontal positioning) and then starts to rotate the elbow as soon as the latter enters its area of attraction (regulating the distance between the anchor point and the tip of the arm). The same gradients can explain the linear motion pattern of an arm tracking dynamically



(a) Marginal gradients for each joint (separate panels) and two levels of recognition density variance Σ_ϕ (green/orange line) computed by backpropagating the error between images with different arm configurations (joint angle: abscissa) and the image of a reference arm configuration (reference angle: red dot).



(b) Gradients for arm (top) and target (bottom) in both Cartesian (left) and joint (right) space.

Figure 16. Visual gradient analysis. $[\theta_0, \theta_1, \theta_2]$ refer to the state of neck/shoulder/elbow joints.

moving targets, when the arm is close to the target (or reference belief): in that case, all gradients provide motion force as explained above.

On the other hand, the gradients relative to the target belief (bottom panels) behave somewhat differently: since this belief is unconstrained and can freely move in the environment, update directions more directly approach the target in all angular coordinates (see the polar plot to the right). Yet, linear belief updates in the polar space still translate to curve directions in the Cartesian space.

DISCUSSION

We presented a normative computational theory based on Active Inference of how neural circuitry in the PPC and DVS supports visually guided goal-directed actions in a dynamically changing environment. Our focus was on the computational basis of encoding dynamic action goals in the PPC through flexible motor intentions. The theory is based on predictive coding (***Doya, 2007***; ***Hohwy, 2013***) and Active Inference (***Friston, 2010***) and evidence suggesting that the PPC performs visuomotor transformations (***Fattori et al., 2017***; ***Galletti and Fattori, 2018***; ***Cisek and Kalaska, 2010***)

and encodes motor plans (***Andersen, 1995***). Accordingly, the PPC is proposed to maintain dynamic expectations of both the current and desired future latent states of both the body and the environment and use them to generate proprioceptive predictions through a hierarchical process at the motor cortex and subcortical structures that descend via efferents down to the motor reflex arcs that ultimately generate movements (***Adams et al., 2013***; ***Versteeg et al., 2021***). On the exteroceptive side, the DVS is proposed to encode a generative model that translates latent-state expectations into visual predictions; discrepancies between sensory-level predictions and actual sensations produce prediction errors that are sent back through the cortical hierarchy to improve the latent-state expectations. The theory unifies research on intention coding in the PPC (***Snyder et al., 1997***) and current views that the PPC estimates the body and the environmental state (***Medendorp and Heed, 2019***), providing specific computational details regarding the involvement of the PPC in goal-directed behavior which extend some perception-bound predictive coding interpretations of the PPC dynamics (***FitzGerald et al., 2015***).

Active Inference unifies predictive coding and motor control through dynamic Bayesian inference; it learns a joint representation of the latent dynamics of the environment and the body to interpret the sensory information and issue control signals through iterative minimization of free energy (***Friston, 2010***, ***2011***; ***Friston et al., 2010***). Variational free energy minimization approximately solves the highly complex sensory interpretation problem that all biological systems face in order to act efficiently, by maintaining a generalized belief over current latent states of the world and its dynamics, and iteratively updating it to accommodate noisy sensory observations. Assuming a multivariate Gaussian form for the generative models trying to predict these observations, inference reduces to the minimization of generalized prediction errors, so that the belief most closely corresponds to both prior expectations and sensory observations. Active Inference further minimizes prediction errors through actions aiming to fulfil prior-biased sensory-level predictions, thus realizing these expectations: indeed, for an agent to restrict itself within a set of preferred states defined by its phenotype and survive, its goal is not only to infer a belief over the causal relationship of the external world, but also to act and sample those states that make the belief true. In this view, an agent controlled by Active Inference does not pursue specific policies by maximizing external rewards as in classical control theories, but by realization of general prior beliefs.

In fact, a goal could be set in Active Inference as a prior over the dynamics function (which specifies the evolution of hidden states over time), either by a sensory-level prediction, which is compared with the actual sensory observation to generate an error driving belief update (***Sancaktar et al., 2020***; ***Oliver et al., 2019***; ***Pio-Lopez et al., 2016***), or directly as a target in the belief space (***Baioumy et al., 2020***). We propose that the PPC implements a more flexible solution, which uses latent-state estimates, that is, belief over the environment, to formalize flexible goals at the latent state-level in terms of a set of intentions and intention functions, each encoding and computing a desired future belief (see Equation 39). A specific intention is selected among other competitive intentions possibly under the control of the PFC and PMd (***Stoianov et al., 2016***) and fulfilled by setting it as a predominant belief trajectory in the generalized form of belief, i.e., as a belief attractor with strong gain (see Equations 41 and 26 and Figure 4). For example, in a typical reaching task, the goal of reaching a specific visual target (among other competitive targets) corresponds to the future expectation that the agent's arm will be over that target; thus, if the agent maintains a belief over the target's position, the corresponding intention links the expected belief over the future body posture with the belief over the currently inferred target, expressed in joint angle coordinates, which thus encodes a specific interaction to realize, depending on object affordances.

We therefore propose that the PPC implements a computational circuitry in which generalized belief encodes trajectories of bodily states and objects in the environment and uses that belief to compute an attractor encoding intentions, e.g., to interact with a particular object. During free energy minimization, the intention is compared with the current belief to produce an intention prediction error driving the agent towards the state it expects to perceive in the future, both mentally, in terms of beliefs over bodily (proprioceptive) state, and physically, through actions fulfilling

proprioceptive predictions that correspond to the desired beliefs. The attractor (intention) could be estimated upon appearance of a target and kept fixed, thus acting as a memory holder - such as the HB in our tests - (***Snyder et al., 1997***; ***Andersen and Cui, 2009***) or it could be dynamically updated (e.g., in case of a moving target), thus implementing flexible intentions. Static and dynamic targets could be set and acted upon under various conditions of visibility (e.g., illumination, sensory noise, temporal occlusion, etc.) and feedback (open vs. closed loop sensorimotor control, ***Todorov (2004)***). Finally, as mentioned above, the PFC could mediate intention-setting and moreover set and manage a series of intentions implementing complex multi-step tasks.

Beyond perceptual accounts (***FitzGerald et al., 2015***), faithful body-state encoding proposals, and flexible sensory-action linking (***Medendorp and Heed, 2019***), our theory postulates that the PPC encodes belief over the body state that is biased towards desired future (intended) states and does not exactly encode its true state. In fact, future-oriented bias in belief - which depends on the magnitude of the intention precisions and gains - is the principal generator of action in continuous Active Inference: in this framework, faithful body state estimate without intention bias will drive no action. Moreover, the theory postulates that the PPC maintains belief over potential environmental targets, and most critically, that the belief over a selected target is dynamically set, or linked as action goal, that is, motor intention. This aspect is close to another view according to which the PPC encodes tasks while bodily states are encoded elsewhere, i.e., in the somatosensory and motor cortices (***Haar and Donchin, 2020***). That latter aspect is inherent in predictive coding under the assumption that these areas encode hierarchical generative models of body state and motor control, which we abstracted altogether with a simple proprioceptive generative model.

In light of these theoretical considerations, the theory predicts several different types of correlates in the PPC related to environment, task, and body states. The former two include correlates of potential spatial targets and of the selected goal positions, that is, intentions, or motor goal, and those indeed have been consistently found in the PPC (***Andersen, 1995***; ***Snyder et al., 1997***; ***Filippini et al., 2018***). The latter includes correlates of intention-biased bodily state estimate which thus are not precise representation of the true state but are still much closer to it than to the intended position (see more on this in section Precision Balance and Conditions for Disorders). Furthermore, the use of generalized belief in Active Inference predicts that the PPC encodes not only static states and errors, but also a detailed estimate of body dynamics, up to a few temporal orders. Indeed, a body of literature report motion-sensitive, or Vision-for-Action activity in the DVP and the PPC (***Galletti and Fattori, 2018***), to which Active Inference gives a detailed computational account. Finally, another set of correlations regards prediction errors. In fact, since in Active Inference belief is estimated through dynamic prediction error minimization, body-environmental transitions involving change of states and tasks will result in transient bursts of activity in error-conveying cells until the error is minimized. Prediction errors conveying upstream information are putatively encoded by pyramidal cortical cells in superficial layers while downstream predictions are encoded by deep pyramidal neurons (***Parr et al., 2022***).

We tested the computational feasibility of the theory on a classical in electrophysiology delayed reaching task in which a monkey is required to reach with its hand a visually presented spatial target, starting movement from a HB (***Breveglieri et al.*** (***2014***), Figure 3b). To do this, we simulated an agent consisting of a coarse 3-DoF limb model and noisy visual and proprioceptive sensors (Figure 3a). Simplified proprioceptive sensors provided a noisy reading of the state of the limb in joint angle coordinates, while visual input was provided by a fixed camera and consisted of an image of the target and the limb. Predictive visual sensory processing simulating the DVS was implemented with a variational autoencoder trained to infer body state and target location, both in the joint angle domain (Figure 3c, 15a). The limbs were animated at the velocity-level with motor control signals computed by the visually guided Active Inference controller.

The computational analysis showed, first and most importantly, that the controller could correctly infer the position of the visual targets (Figure 5a-c), use it to compute and set motor goals in terms of prior beliefs on the future body state through intention functions, and perform adequate

and smooth reach movements (Figure 5 d-f) under both closed and open loop conditions, that is, with and without visual feedback of current posture during reach movement (Figure 6).

We then systematically investigated the effect of noise (or precision) on various functional components in the current implementation (Figure 8), starting with the balance of self-perception precision between proprioceptive and visual sensory model: a noiseless Active Inference agent (BL condition) resulted in the best performance, with a stable final approach and accuracy only limited by the quality of visual target estimation. Among the noisy conditions, pure proprioceptive control resulted in the lowest performance, as expected. Motor control driven by both proprioceptive and visual feedback with balanced precision between the two domains resulted in improved reach accuracy and greatly improved arm belief stability (Figure 9). The effect on accuracy was mainly due to the inclusion of visual information in belief inference, but also to the slower updates of motor control signals due to the decreased confidence in proprioceptive input.

The increased stability of arm belief did not improve movement stability as increasing confidence in visual input also increased the discrepancy between belief and action update, since the latter only relies on noisy proprioceptive observations. In fact, we showed that if we remove the plausibility constraint that motor control is driven only by proprioceptive predictions and thus let actions minimize prediction errors from all sensory domains, reach performance and arm belief stability greatly increase (Figure 10). In particular, any combination of visual and proprioceptive feedback improved performance relative to control driven by feedback from a single sensory domain.

On the other hand, the instability due to the difference in update directions between belief and action could be balanced by other mechanisms that we have not considered here. For example, we assumed that the same pathway is used for both control and belief inference, but it seems that the motor cortex generates different predictions depending on the brain areas which it interacts with: purely proprioceptive predictions for motor control, whose prediction error is suppressed at the lowest level of the hierarchy, and rich somatosensory predictions for latent state inference, which integrates somatic sensations at different hierarchical levels (***Adams et al., 2013***).

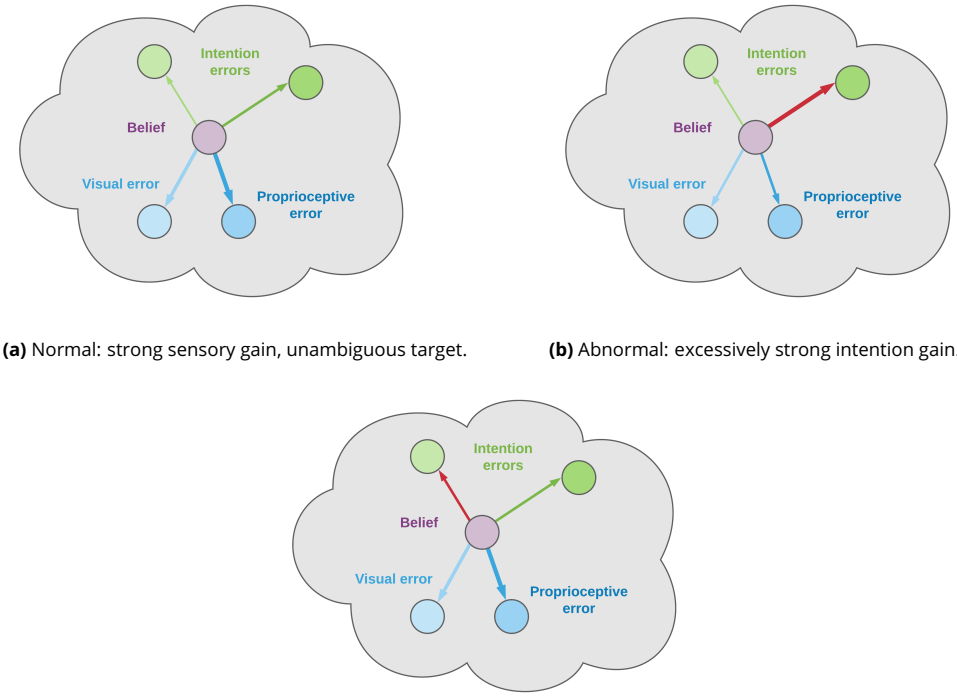
Intention precisions or attractor gains affected performance, too. First, they affected reach time: as expected, the greater the gain, the faster the movement. However, fast movements come at a cost: increased gains generally resulted in less precise movements and decreased stability during the final reach period. Finally, higher visual target precisions decreased perception time and improved perception accuracy, but decreased perception stability.

We also investigated the effects of movement onset policy: response delay allows investigating perceptual and motor preparatory processes separately from motor control and action execution. We found that delayed response decreased movement time with respect to a policy that requires immediate response (Figure 11), which fits the behavioral pattern (***Shenoy et al., 2013***). Apparently, this is due to the need of estimating, in the latter condition, the target position "on the fly", and constantly adapting the intention (body-state attractor) according to the updated estimate of the target belief. The advantage of allowing some preparatory time becomes clear in an anecdotal fly-catching task: this will result in faster movement, thus, increased chances of success. This comes with the critical contribution of PPC neurons which systematically modulate their activity during the preparatory period (***Shenoy et al., 2013***), and here provided specific predictions for the computations performed in the PPC.

Notably, the immediate-response policy allowed the Active Inference controller to perform actions under dynamic environmental conditions, such as tracking moving objects. Free energy minimization resulted in rapid target detection also in this case, and maintained subthreshold perception error on moving targets (Figure 12), that allowed precise target tracking after an initial target-reach period. Thus, the agent was able to maintain correct belief estimates over both a dynamic target and body posture (the arm configuration) and dynamically update the intentions according to the current target position.

Free energy minimization over generalized belief is at the heart of the performance of goal-

directed Active Inference. That is, achieving dynamic goals (here, reaching a visual target) was possible by dynamically setting continuously changing attractors derived from the current belief on the latent state of the environment (i.e., target configuration) and minimizing the difference between the attractors and the belief. In fact, we showed that the agent was able to quickly infer and maintain the instantaneous trajectory of the latent states, under both static and dynamic target conditions (Figure 14).



(a) Normal: strong sensory gain, unambiguous target.

(b) Abnormal: excessively strong intention gain.

(c) Abnormal: ambiguous target.

Figure 17. Normal and abnormal intention gains and precisions. Arrow width represents strength of prediction errors. (a) In normal conditions, the sensory contribution to free energy minimization is bigger than that of the intentions and there is a clear intention to fulfil. (b) Abnormally strong intention gain k drives belief away from the true joint state. (c) Too-close precision levels γ hinder implementing any competing intention.

Precision Balance and Conditions for Disorders

Based on our computational analysis, it becomes clear that some motor and behavioral disorders could be due to the lack of proper sensory and intention precisions (***Adams et al., 2021***). Here, we illustrate the normal condition and two types of potentially improper precision balance that could become a causal condition for neurological disorders.

To begin with, Figure 17a illustrates the condition for normal functioning, which is such that the contribution of a single intention to belief update (which, as reminder to the reader, is proportional to the gain of intentions k) is sufficiently small with respect to the sensory contribution. In this case, during free energy minimization, the system dynamics smoothly moves the belief towards the strongest goal, along with precise tracking of the true latent state and sensory signal of the limbs, allowing thus to compute correct motor control errors and perform smooth action execution.

A critical abnormal condition is one in which the intention gain k is too strong, as illustrated in Figure 17b. In this case, the belief moves too rapidly towards the goal without being able to match the proprioceptive observations, which result in computing incorrect motor control signals. Another abnormal condition is caused by too close precisions γ of competitive intentions, which is

likely to result in opposing belief updates and thus prevent the fulfillment of any of the competing goals (as in Figure 17c). This situation might manifest itself in terms of motor onset failure or oscillatory behavior.

Limitations and Future Directions

Our focus here was on lower-level intention coding in the PPC, which directly deals with motor plans and motor control. Further elaborations will extend the theory with higher-level aspects of cognitive control, including intention structuring (dealt by PM), phasing (SMA) (***Gallego et al., 2022***), planning, and goal selection (HC, PFC) (***Stoianov et al., 2016***; ***Pezzulo et al., 2019***).

Motor control operated here in an inner belief space belonging to the joint angle domain, which is generally suboptimal in the external Cartesian space. In fact, we assumed that neural activity in the PPC encodes generalized beliefs over targets and body in a motor-related domain; however, neural data suggest that neurons in the motor cortex encode motor trajectories also in extrinsic coordinates (***Adams et al., 2013***; ***Cohen and Andersen, 2002***), and a more realistic model should include representations encoding states in both intrinsic and extrinsic frames of reference. A functional correlate of the motor cortex should represent future states - which were defined here implicitly in the intention prediction errors and dynamics functions - and transform desired trajectories from Cartesian coordinates to proprioceptive predictions in intrinsic state-space. This transformation is different from optimal control planning since the optimization of a classical inverse model reduces to a more manageable inference problem. Finally, although free energy minimization in a single belief domain is faster and more parsimonious, more precise visually guided motor control should include continuous activation of visual hand-target attractors at a certain "neural energy" cost, since motor areas interact with and generate predictions in multiple sensory domains (e.g., somatosensory).

The current implementation just partially covers the three main domains of sensorimotor learning through fitting a predictive forward control; for example, it does not fully include reactive, stimuli-driven control such as obstacle avoidance, although we show that it can perform static and dynamic tasks (target reaching/tracking).

Since our focus was on the theoretical introduction of intentionality in Active Inference, perceptual inference, visually guided sensorimotor control, and motor behavior of the specific implementation were only partially characterized in a simple reaching task. However, fundamental properties of the physical model, including geometry, mass, and friction, strongly influence the resulting motion dynamics, and thus the entire inferential process. This implementation does not adopt other important neural and biomechanical specificities such as signal delay and joint friction (***Wolpert and Flanagan, 2016***); also, it only includes proprioceptive and visual feedback. However, it could be easily extended to accommodate other sensory modalities (e.g., tactile sensations) with a rich generative model such as the visual one here implemented. Therefore, further planned computational analysis will use a more realistic physical arm model, additional actuators, extended belief space, and extend the complexity of the intention functions in order to investigate the capacity of the theory to explain in-depth neural-level, cognitive, and kinematic phenomena related to motor learning, motion perception, motor planning, and so on.

Although we considered Active Inference with just a single layer of intentions, the structure represented in Figure 2 could be scaled hierarchically by considering intermediate goals between high-level intentions and low-level sensory generative models, for example, by combining discrete and continuous Active Inference for planning and movement execution (***Friston et al., 2017a***,b; ***Parr et al., 2020***; ***Sajid et al., 2021***). According to the free energy principle, the agent will then choose goals and subgoals and rely on specific sensory modalities such that free energy is minimized at every level of the hierarchy based on the prediction error coming from the level below. This hierarchical formalization will provide an explicit basis for motor planning, including tasks such as object manipulation. Indeed, although the current implementation performs well on spatial tasks such as reaching in a dynamically changing environment, it cannot implement composite goals which the

brain nevertheless needs to handle, where the current model will fail to provide a correct direction of belief and action update. Therefore, the desired belief should be sufficiently simple, e.g., consisting of plain spatial targets, in order for the sensory generative models to find a correct gradient (indeed, all the hard work of motor planning is essentially assigned to them). On the other hand, an agent with an enhanced hierarchical structure of belief that has the possibility to encode higher level intentions in a discrete domain and infer policies based on the *expected free energy* will be able to dynamically modify its behavior in order to react to environmental changes. An extended implementation of this kind will be an interesting and more biologically plausible alternative to standard reinforcement learning (***Stoianov et al., 2018***).

ACKNOWLEDGEMENTS

This research received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under H2020-EIC-FETPROACT-2019 Grant Agreement 951910 (MAIA) to IS, Grant Agreement No 945539 (Human Brain Project SGA3), the European Research Council under Grant Agreement No. 820213 (ThinkAhead), and from the Italian Ministry for Research MIUR under Grant Agreement PRIN 2017KZNZLN (PACE) to IS.

References

Adams RA, Shipp S, Friston KJ. Predictions not commands: Active inference in the motor system. Brain Structure and Function. 2013; 218(3):611–643. doi: 10.1007/s00429-012-0475-5.

Adams RA, Vincent P, Benrimoh D, Friston KJ, Parr T. Everything is connected: Inference and attractors in delusions. Schizophrenia Research. 2021; (March). doi: 10.1016/j.schres.2021.07.032.

Andersen RA. Encoding of intention and spatial location in the posterior parietal cortex. Cerebral Cortex. 1995; 5(5):457–469. doi: 10.1093/cercor/5.5.457.

Andersen RA, Cui H. Intention, Action Planning, and Decision Making in Parietal-Frontal Circuits. Neuron. 2009; 63(5):568–583. http://dx.doi.org/10.1016/j.neuron.2009.08.028, doi: 10.1016/j.neuron.2009.08.028.

Baioumy M, Duckworth P, Lacerda B, Hawes N. Active inference for integrated state-estimation, control, and learning. arXiv. 2020; .

Baltieri M, Buckley CL. PID control as a process of active inference with linear generative models. Entropy. 2019; 21(3). doi: 10.3390/e21030257.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. Canonical Microcircuits for Predictive Coding. Neuron. 2012; 76(4):695–711. http://dx.doi.org/10.1016/j.neuron.2012.10.038, doi: 10.1016/j.neuron.2012.10.038.

Bishop CM. Pattern Recognition & Machine Learning. New York, New York, USA: Springer; 2006.

Bogacz R. A tutorial on the free-energy framework for modelling perception and learning. Journal of Mathematical Psychology. 2017; 76:198–211. doi: 10.1016/j.jmp.2015.11.003.

Breveglieri R, Galletti C, Dal Bò G, Hadjidimitrakis K, Fattori P. Multiple Aspects of Neural Activity during Reaching Preparation in the Medial Posterior Parietal Area V6A. Journal of cognitive neuroscience. 2014; 26(4):879–895. doi: 10.1162/jocn_a_00510.

Buckley CL, Kim CS, McGregor S, Seth AK. The free energy principle for action and perception: A mathematical review. Journal of Mathematical Psychology. 2017; 81:55–79. doi: 10.1016/j.jmp.2017.09.004.

Cisek P, Kalaska JF. Neural Mechanisms for Interacting with a World Full of Action Choices. Annual Review of Neuroscience. 2010; 33:269–298. doi: 10.1146/annurev.neuro.051508.135409.

Cohen YE, Andersen RA. A common reference frame for movement plans in the posterior parietal cortex. Nature Reviews Neuroscience. 2002; 3(7):553–562. doi: 10.1038/nrn873.

Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews Neuroscience. 2002; 3(3):201–215. doi: 10.1038/nrn755.

Desmurget M, Epstein CM, Turner RS, Prablanc C, Alexander GE, Grafton ST. PPC and visually directing reaching to targets. Nature Ne. 1999; 2(6):563–567. doi: doi.org/10.1038/9219.

Doya K. Bayesian Brain: Probabilistic Approaches to Neural Coding. Press M, editor; 2007.

Fattori P, Breveglieri R, Bosco A, Gamberini M, Galletti C. Vision for Prehension in the Medial Parietal Cortex. Cerebral Cortex. 2017; 27(February):1149–1163. http://cercor.oxfordjournals.org/cgi/doi/10.1093/cercor/bhv302, doi: 10.1093/cercor/bhv302.

Filippini M, Breveglieri R, Ali Akhras M, Bosco A, Chinellato E, Fattori P. Decoding information for grasping from the macaque dorsomedial visual stream. Journal of Neuroscience. 2017; 37(16):4311–4322. doi: 10.1523/JNEUROSCI.3077-16.2017.

Filippini M, Breveglieri R, Hadjidimitrakis K, Bosco A, Fattori P. Prediction of Reach Goals in Depth and Direction from the Parietal Cortex. Cell Reports. 2018; 23(3):725–732. https://doi.org/10.1016/j.celrep.2018.03.090, doi: 10.1016/j.celrep.2018.03.090.

FitzGerald THB, Moran RJ, Friston KJ, Dolan RJ. Precision and neuronal dynamics in the human posterior parietal cortex during evidence accumulation. NeuroImage. 2015; 107:219–228. doi: 10.1016/j.neuroimage.2014.12.015.

Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F, Rizzolotti G. Parietal lobe: From action organization to intention understanding. Science. 2005; 308(5722):662–667. doi: 10.1126/science.1106138.

Friston KJ, Trujillo-Barreto N, Daunizeau J. DEM: A variational treatment of dynamic systems. NeuroImage. 2008; 41(3):849–885. doi: 10.1016/j.neuroimage.2008.02.054.

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

Friston K. The free-energy principle: A unified brain theory? Nature Reviews Neuroscience. 2010; 11(2):127–138. http://dx.doi.org/10.1038/nrn2787, doi: 10.1038/nrn2787.

Friston K. What is optimal about motor control? Neuron. 2011; 72(3):488–498. http://dx.doi.org/10.1016/j.neuron.2011.10.018, doi: 10.1016/j.neuron.2011.10.018.

Friston K. The history of the future of the Bayesian brain. NeuroImage. 2012; 62(2):1230–1233. http://dx.doi.org/10.1016/j.neuroimage.2011.10.004, doi: 10.1016/j.neuroimage.2011.10.004.

Friston K, Kiebel S. Predictive coding under the free-energy principle. Philosophical Transactions of the Royal Society B: Biological Sciences. 2009; 364(1521):1211–1221. doi: 10.1098/rstb.2008.0300.

Friston K, Mattout J, Trujillo-Barreto N, Ashburner J, Penny W. Variational free energy and the Laplace approximation. NeuroImage. 2007; 34(1):220–234. doi: 10.1016/j.neuroimage.2006.08.035.

Friston KJ. Functional integration and inference in the brain. Progress in neurobiology. 2002 oct; 68(2):113–43. http://www.ncbi.nlm.nih.gov/pubmed/12450490.

Friston KJ. A theory of cortical responses. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2005 apr; 360(1456):815–36. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1569488{&}tool=pmcentrez{&}rendertype=abstract, doi: 10.1098/rstb.2005.1622.

Friston KJ, Daunizeau J, Kiebel SJ. Reinforcement learning or active inference? PLoS ONE. 2009; 4(7). doi: 10.1371/journal.pone.0006421.

Friston KJ, Daunizeau J, Kilner J, Kiebel SJ. Action and behavior: A free-energy formulation. Biological Cybernetics. 2010; 102(3):227–260. doi: 10.1007/s00422-010-0364-z.

Friston KJ, Mattout J, Kilner J. Action understanding and active inference. Biological cybernetics. 2011 feb; 104(1-2):137–60. doi: 10.1007/s00422-011-0424-z.

Friston KJ, Parr T, de Vries B. The graphical brain: Belief propagation and active inference. . 2017; 1(4):381–414. http://dx.doi.org/10.1162/NETN_a_00018, doi: 10.1162/NETN.

Friston KJ, Rosch R, Parr T, Price C, Bowman H. Deep temporal models and active inference. Neuroscience and Biobehavioral Reviews. 2017; 77(November 2016):388–402. http://dx.doi.org/10.1016/j.neubiorev.2017.04.009, doi: 10.1016/j.neubiorev.2017.04.009.

Friston KJ, Samothrakis S, Montague R. Active inference and agency: optimal control without cost functions. Biological cybernetics. 2012; (106):523–541. doi: 10.1007/s00422-012-0512-8.

Gallego JA, Makin TR, McDougle SD. Going beyond primary motor cortex to improve brain–computer interfaces. Trends in Neurosciences. 2022; 45(3):176–183. https://doi.org/10.1016/j.tins.2021.12.006, doi: 10.1016/j.tins.2021.12.006.

Galletti C, Fattori P. The dorsal visual stream revisited: Stable circuits or dynamic pathways? Cortex. 2018; 98:203–217. https://doi.org/10.1016/j.cortex.2017.01.009, doi: 10.1016/j.cortex.2017.01.009.

Gamberini M, Passarelli L, Filippini M, Fattori P, Galletti C. Vision for action: thalamic and cortical inputs to the macaque superior parietal lobule. Brain Structure and Function. 2021; 226(9):2951–2966. https://doi.org/10.1007/s00429-021-02377-7, doi: 10.1007/s00429-021-02377-7.

Genovesio A, Tsujimoto S, Wise SP. Encoding goals but not abstract magnitude in the primate prefrontal cortex. Neuron. 2012 may; 74(4):656–62. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3400847&tool=pmcentrez&rendertype=abstract, doi: 10.1016/j.neuron.2012.02.023.

Goodfellow IJ, Bengio Y, Courville A. Deep Learning. Cambridge, MA, USA: MIT Press; 2016. http://www.deeplearningbook.org.

Haar S, Donchin O. A revised computational neuroanatomy for motor control. Journal of Cognitive Neuroscience. 2020; 32(10):1823–1836. doi: 10.1162/jocn_a_01602.

Hohwy J. The Predictive Mind. Oxford University Press UK; 2013.

Kaplan R, Friston KJ. Planning and navigation as active inference. Biological Cybernetics. 2018; 112(4):323–343. https://doi.org/10.1007/s00422-018-0753-2, doi: 10.1007/s00422-018-0753-2.

Keele SW, Posner MI. Processing of Visual Feedback in Rapid Movements. Journal of Experimental Psychology. 1968; 77(1):155–158. doi: 10.1037/h0025754.

Kikuchi Y, Hamada Y. Geometric characters of the radius and tibia in Macaca mulatta and Macaca fascicularis. Primates. 2009; 50(2):169–183. doi: 10.1007/s10329-008-0120-3.

Kingma DP, Welling M. Auto-encoding variational bayes. 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings. 2014; p. 1–14.

Lanillos P, Cheng G. Adaptive Robot Body Learning and Estimation Through Predictive Coding. IEEE International Conference on Intelligent Robots and Systems. 2018; p. 4083–4090. doi: 10.1109/IROS.2018.8593684.

Lau HC, Rogers RD, Haggard P, Passingham RE. Attention to Intention. Sicence. 2004; 303(5661):1208–1210. doi: 10.1126/science.1090973.

Levine S. Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review. . 2018; http://arxiv.org/abs/1805.00909.

Limanowski J, Friston K. Active inference under visuo-proprioceptive conflict: Simulation and empirical results. Scientific Reports. 2020; 10(1):1–14. http://dx.doi.org/10.1038/s41598-020-61097-w, doi: 10.1038/s41598-020-61097-w.

Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. Nature Neuroscience. 2006; 9(11):1432–1438. doi: 10.1038/nn1790.

Medendorp WP, Heed T. State estimation in posterior parietal cortex: Distinct poles of environmental and bodily states. Progress in Neurobiology. 2019; 183(April):101691. doi: 10.1016/j.pneurobio.2019.101691.

Millidge B, Tschantz A, Seth AK, Buckley CL. On the relationship between active inference and control as inference. Communications in Computer and Information Science. 2020; 1326:3–11. doi: 10.1007/978-3-030-64919-7_1.

Oliver G, Lanillos P, Cheng G. Active inference body perception and action for humanoid robots. . 2019; http://arxiv.org/abs/1906.03022.

Parr T, Friston KJ. The anatomy of inference: Generative models and brain structure. Frontiers in Computational Neuroscience. 2018; 12(November). doi: 10.3389/fncom.2018.00090.

Parr T, Pezzulo G, Friston KJ. Active inference: the free energy principle in mind, brain, and behavior. Cambridge, MA: MIT Press; 2021.

Parr T, Pezzulo G, Friston KJ. Active Inference: The Free Energy Principle in Mind, Brain, and Behavior. The MIT Press; 2022. https://doi.org/10.7551/mitpress/12441.001.0001, doi: 10.7551/mitpress/12441.001.0001.

Parr T, Rikhye RV, Halassa MM, Friston KJ. Prefrontal Computation as Active Inference. Cerebral Cortex. 2020; 30(2):682–695. doi: 10.1093/cercor/bhz118.

Pezzulo G, Cisek P. Navigating the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition. Trends in Cognitive Sciences. 2016; 20(6):414–424. http://dx.doi.org/10.1016/j.tics.2016.03.013, doi: 10.1016/j.tics.2016.03.013.

Pezzulo G, Donnarumma F, Dindo H, D'Ausilio A, Konvalinka I, Castelfranchi C. The body talks: Sensorimotor communication and its brain and kinematic signatures. Physics of Life Reviews. 2019; 28:1–21. https://doi.org/10.1016/j.plrev.2018.06.014, doi: 10.1016/j.plrev.2018.06.014.

Pezzulo G, Donnarumma F, Iodice P, Maisto D, Stoianov I. Model-based approaches to active perception and control. Entropy. 2017; 19(6). doi: 10.3390/e19060266.

Pezzulo G, Rigoli F, Friston KJ. Hierarchical Active Inference: A Theory of Motivated Control. Trends in Cognitive Sciences. 2018; 22(4):294–306. http://dx.doi.org/10.1016/j.tics.2018.01.009, doi: 10.1016/j.tics.2018.01.009.

Pio-Lopez L, Nizard A, Friston K, Pezzulo G. Active inference and robot control: A case study. Journal of the Royal Society Interface. 2016; 13(122). doi: 10.1098/rsif.2016.0616.

Rao RPN, Ballard DH. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience. 1999; 2(1):79–87. doi: 10.1038/4580.

Rood T, van Gerven M, Lanillos P. A deep active inference model of the rubber-hand illusion. . 2020; http://arxiv.org/abs/2008.07408.

Sajid N, Ball PJ, Parr T, Friston KJ. Active inference: demystified and compared. Neural Computation. 2021; 33(3):674–712. doi: 10.1162/neco_a_01357.

Sancaktar C, van Gerven MAJ, Lanillos P. End-to-End Pixel-Based Deep Active Inference for Body Perception and Action. In: *2020 Joint IEEE 10th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*; 2020. p. 1–8. doi: 10.1109/icdl-epirob48136.2020.9278105.

Saunders JA, Knill DC. Humans use continuous visual feedback from the hand to control fast reaching movements. Experimental Brain Research. 2003; 152(3):341–352. doi: 10.1007/s00221-003-1525-2.

Shadmehr R, Krakauer JW. A computational neuroanatomy for motor control. Experimental Brain Research. 2008; 185(3):359–381. doi: 10.1007/s00221-008-1280-5.

Shenoy KV, Sahani M, Churchland MM. Cortical Control of Arm Movements: A Dynamical Systems Perspective. Annual Review of Neuroscience. 2013; 36(1):337–359. doi: 10.1146/annurev-neuro-062111-150509.

Snyder LH, Batista AP, Andersen RA. Coding of intention in the posterior parietal cortex. Nature. 1997; 386(6621):167–170. doi: 10.1038/386167a0.

Snyder LH, Batista AP, Andersen RA. Intention-related activity in the posterior parietal cortex: A review. Vision Research. 2000; 40(10-12):1433–1441. doi: 10.1016/S0042-6989(00)00052-3.

Srinivasan SS, Gutierrez-Arango S, Teng ACE, Israel E, Song H, Bailey ZK, Carty MJ, Freed LE, Herr HM. Neural interfacing architecture enables enhanced motor control and residual limb functionality postamputation. Proceedings of the National Academy of Sciences of the United States of America. 2021; 118(9). doi: 10.1073/pnas.2019555118.

Stoianov I, Pennartz C, Lansink C, Pezzulo G. Model-based spatial navigation in the hippocampus-ventral striatum circuit: a computational analysis. Plos Computational Biology. 2018; 14(9):1–28. doi: 10.1371/journal.pcbi.1006316.

Stoianov I, Genovesio A, Pezzulo G. Prefrontal goal codes emerge as latent states in probabilistic value learning. Journal of Cognitive Neuroscience. 2016; 28(1):140–157. http://www.mitpressjournals.org/doi/10.1162/jocn_a_00886, doi: 10.1162/jocn_a_00886.

Todorov E. Optimality principles in sensorimotor control. Nature Neuroscience. 2004; 7:907–915. doi: 10.1038/nn1309.

Toussaint M, Storkey A. Probabilistic inference for solving discrete and continuous state Markov Decision Processes. ACM International Conference Proceeding Series. 2006; 148:945–952. doi: 10.1145/1143844.1143963.

Tuthill JC, Azim E. Proprioception. Current Biology. 2018; 28(5):R194–R203. doi: 10.1016/j.cub.2018.01.064.

Velliste M, Perel S, Spalding MC, Whitford AS, Schwartz AB. Cortical control of a prosthetic arm for self-feeding. Nature. 2008; 453(7198):1098–1101. doi: 10.1038/nature06996.

Versteeg C, Rosenow JM, Bensmaia SJ, Miller LE. Encoding of limb state by single neurons in the cuneate nucleus of awake monkeys. Journal of Neurophysiology. 2021; 126(2):693–706. doi: 10.1152/jn.00568.2020.

Wolpert DM, Flanagan JR. Computations underlying sensorimotor learning. Current Opinion in Neurobiology. 2016; 37:7–11. http://dx.doi.org/10.1016/j.conb.2015.12.003, doi: 10.1016/j.conb.2015.12.003.

SUPPLEMENTARY MATERIAL

Parameter	Value
trunk_length	17
trunk_width	16
trunk_limits	(0, 10)
shoulder_length	27
shoulder_width	14
shoulder_limits	(0, 130)
elbow_length	38
elbow_width	12
elbow_limits	(0, 130)

Table 1. Arm model parameters.

Parameter	Value
n_datapoints	20000
n_epochs	100
n_batch	32
learning_rate	0.001
variance	0.02

Table 2. Variational autoencoder parameters.

Parameter	Value
Δ_t	0.6
α	0.4
π_p	0.5
π_{v_a}	2e-5
π_{v_t}	4e-4
π_{v_h}	1.0
β	0.1
k	0.1
w_p	2e-3
w_a	0.0
n_steps	300
n_trials	1000

Table 3. Active Inference parameters.