
DYNAMIC PLANNING IN HIERARCHICAL ACTIVE INFERENCE

Matteo Priorelli

Institute of Cognitive Sciences and Technologies
National Research Council of Italy, Padova
Sapienza University of Rome, Italy
matteo.priorelli@gmail.com

Ivilin Peev Stoianov

Institute of Cognitive Sciences and Technologies
National Research Council of Italy
Padova, Italy
ivilinpeev.stoianov@cnr.it

ABSTRACT

By dynamic planning, we refer to the ability of the human brain to infer and impose motor trajectories related to cognitive decisions. A recent paradigm, *active inference*, brings fundamental insights into the adaptation of biological organisms, constantly striving to minimize prediction errors to restrict themselves to life-compatible states. Over the past years, many studies have shown how human and animal behaviors could be explained in terms of active inference – either as discrete decision-making or continuous motor control – inspiring innovative solutions in robotics and artificial intelligence. Still, the literature lacks a comprehensive outlook on effectively planning realistic actions in changing environments. Setting ourselves the goal of modeling complex tasks such as tool use, we delve into the topic of dynamic planning in active inference, keeping in mind two crucial aspects of biological behavior: the capacity to understand and exploit affordances for object manipulation, and to learn the hierarchical interactions between the self and the environment, including other agents. We start from a simple unit and gradually describe more advanced structures, comparing recently proposed design choices and providing basic examples. This study distances itself from traditional views centered on neural networks and reinforcement learning, and points toward a yet unexplored direction in active inference: hybrid representations in hierarchical models.

1 Introduction

Three characteristics of the brain are relevant to tasks involving planning in changing environments, such as tool use. First, the brain’s ability to maintain estimates not only of bodily states, but also of external physical variables in relation to the self. Studies have shown how the Posterior Parietal Cortex (PPC) of the monkey brain encodes objects with sensorimotor representations reflecting the body structure [1, 2]. These representations are extremely useful for object manipulation since they account efficiently for the action possibilities provided by the object, also known as *affordances* [3]. For instance, one might encode a cup in different ways (power versus precision grips) based on whether one wants to throw it or drink from it. Further, to act timely in a dynamic environment, the PPC can encode multiple objects in parallel during sequences of actions, even when there is a considerable delay between different subgoals [4].

A second characteristic regards flexible and deep hierarchies. Hierarchical structures are so pervasive that they not only exist as causal relationships between physical properties of the environment, but are also inherent to how biological organisms act over it. Even the most complex kinematic structures of animals follow a rigid hierarchical strategy, whereby different limbs propagate from a body-centered reference frame. The hierarchical modularity of brain functional networks is widely recognized [5, 6], as well as the representation of the body schema in somatosensory and motor areas [7], and the organization of hierarchical motor sequences concerning parietal and premotor cortices [8]. In particular, the body schema is not a static entity but changes in concurrence to the development of the human body during childhood and adolescence [9]. Surprisingly, the nervous system is able to relate external objects to the self in a way that, although not reflecting the actual causal relationships between the body and the environment, is the most suitable for better operating in a specific context. Physiological studies have demonstrated that, with extensive tool use, parietal and motor areas of the monkey brain gradually adapt to make room for the tool, increasing the length of the perceived limb [10, 11]. This adaptation is highly plastic, assimilating objects in a very short time [12] and inducing altered somatosensory representations of the body morphology that persist even after tool use [13].

A third characteristic is the ability to construct a dynamic discretized plan based on continuous sensory evidence. Complex tasks involve decision-making, which the brain is known to realize via several methods [14]. Among them, one is particularly relevant: planning for deliberation, also known as *vicarious trials and errors*, whereby an action is selected after several alternatives have been generated and evaluated [15]. One of the most intriguing aspects of human planning is the capacity to imagine, or endogenously generate dynamic representations of future states, including potential trajectories and subgoals that bring to such states [16, 17]. The hippocampus is a key neural structure known to support trajectory generation, although planning is accomplished in concert with other areas [15].

How does the human brain capture the hierarchical organization and dynamics of the self and the environment to afford purposeful planning? One recent theory is that of *predictive coding* [18, 19, 20, 21], which has been attracting increasing interest in recent years and proposes itself as a unifying paradigm of cortical function. According to predictive coding, living beings make sense of the world by building an internal generative model that imitates the causal relationships of the external generative process. From a high-level hypothesis about the world, a cascade of neural predictions takes place, eventually leading to a low-level guess about sensory evidence. Comparing the model's guess with the sensorium triggers another cascade of *prediction errors* that travel back to the deepest cortical levels. The model iteratively refines its structure until all the prediction errors are minimized, that is, until it correctly predicts what will happen next. This optimization differs from the more traditional view of deep learning, in that the message passing is local and what climbs up the hierarchy does not signal the detection of a feature, but how much the model is *surprised* about its prediction. Besides having stimulated cognitive and neural studies under several circumstances [22, 23, 24, 25], this theory has also influenced novel directions in machine learning: Predictive Coding Networks (PCNs) have been shown to generalize well to classification or regression tasks [26, 27], with key advantages compared to neural networks and still approximating the backpropagation algorithm [28, 29, 30].

While predictive coding can elucidate illusions and visual phenomena such as binocular rivalry [31], it explains just the first (perceptual) half of the story. More specifically, it does not explain why interactions with the environment occur – a process that results, considering the above example, in the monkey brain actively distorting its body schema during tool use. On this trail, a second innovative perspective has been proposed, aspiring to unveil a unified first principle not just on cortical function, but on the behavior of all living organisms. This perspective, called *active inference* [32, 33, 34, 35], is grounded on the same theoretical basis of predictive coding but further assumes two key aspects of biological behavior. First, that a living being does not maintain static hypotheses about the world but also constructs internal dynamics – either as instantaneous trajectories or future states – to anticipate the unfolding of events occurring at different timescales. Second, that these dynamic hypotheses can be fulfilled by movements. The latter assumption replaces models with *agents*, conveying a somewhat counterintuitive but insightful implication: while perception lets the agent's hypothesis conform to the environment (as in predictive coding), action forces the environment to conform to the hypothesis – by sampling those observations that make the hypothesis true. If such hypotheses or *beliefs* correspond to desired states defined, e.g., by the phenotype, cycling between action and perception ultimately allows the agent to survive. This is the core of the so-called *free energy principle*, which states that in order to maintain homeostasis, all organisms must constantly and actively minimize the difference between their sensory states and their expectations based on a small set of life-compatible options. Giving a practical example, if I believe to find myself with a tool in hand, I will try with all my strength to observe visual images of the tool in my hand; in doing this, a combined reaching and grasping action happens. This view distances itself from the stimulus-response mapping widely established in neuroscience, and evidence indicates that it could be more biologically plausible than optimal control and Reinforcement Learning (RL) [36, 37, 38, 39].

Active inference implementations can be divided into two frameworks, which have been used to simulate human and animal behaviors under the two complementary aspects of motor control [40, 41, 42, 43, 44, 45, 46] and decision-making [47, 48, 49, 50, 51]. In principle, active inference might be key for understanding how goal-directed behavior emerges in the human brain [52]. For instance, relevant objects used for manipulation may gradually become part of one's identity through a closed loop between motor commands and sensory evidence, meaning that the boundary of the self from the environment increases whenever the agent predicts the consequences of its own movements [53]. Additionally, active inference might lead to key advances with current artificial agents, taking forward a promising research area known as *planning as inference* [54, 55, 56]. The three characteristics delineated above are fundamental to designing active inference agents that can tackle real-life applications such as tool use. But how to combine them into a single view? In other words, how to perform dynamic planning with hierarchical structures of several objects?

To answer this question, in this study we explore an alternative direction in active inference, i.e., toward hybrid computations in hierarchical systems. We analyze many design choices that have been applied in the motor control domain, with an in-depth look at object affordances, deep hierarchies, and planning with continuous signals. Asking ourselves how to model tool use, we start from a simple unit and construct richer modules that can be linked in a hierarchical fashion, exhibiting interesting high-level features. In Chapter 2, we consider a single-DoF agent and explore how to account for affordances and realize a multi-step behavior in continuous time only. In Chapter 3, we analyze the implications of combining different units in a single network, using more complex kinematic configurations and distinguishing between intrinsic and extrinsic dynamics. In Chapter 4, we describe the advantages of using discrete decision-making in continuous environments, focusing on hybrid structures and drawing some parallelisms between the two worlds. Finally, in the Discussion we elaborate on the benefits of addressing discrete and continuous representations together, and give a few suggestions for future work on this subject.

2 Modeling affordances

In this chapter, we explain the inference mechanisms of a basic unit in continuous time. We then discuss one by one the changes and features that we introduce, in order to achieve a multi-step behavior in simple tasks that do not require deep hierarchical modeling nor online replanning.

The continuous-time active inference framework [57, 36, 40] – generally compared to the low-level sensorimotor loops – makes use of generalized filtering [58] to model instantaneous trajectories of the self and the environment; these trajectories are inferred by minimization of a quantity called *variational free energy*, which is the negative of what in machine learning is known as the *evidence lower bound*. Differently from optimal control, motor commands in active inference derive from proprioceptive predictions that are fulfilled by classical spinal reflex arcs [38]. This eliminates the need for cost functions – as the inverse model maps from proprioceptive (and not latent) states to actions – and replaces a control problem with an inference problem [37].

Modeling of objects in active inference has been recently done in the context of active object reconstruction [59, 60, 61, 62] – where an agent encoded independent representations for multiple elements, and used action to more accurately infer its dynamics; for simulating oculomotor behavior [63] – where the dynamics of a target belief was biased by a hidden location; or for analyzing epistemic affordance [51], i.e., the changes in affordance of different objects in relation to the agent’s beliefs. In continuous time, such affordances can be expressed in intrinsic reference frames corresponding to potential agent’s configurations, defining specific ways to interact with the objects. Manipulating these additional beliefs depending on the agent’s intentions [64] permits effectively operating in dynamic contexts, e.g., tracking a target with the eyes [63], or grasping an object on the fly [65] and placing it at a goal position [66].

2.1 A simple agent

The most elementary unit is represented in Figure 1a. This is the simplest formulation of a continuous-time active inference agent, where we kept only the key nodes. This allows us to easily describe a velocity-controlled dynamic system with the following likelihood \mathbf{g}_p and dynamics \mathbf{f} :

$$\begin{aligned} \mathbf{o}_p &= \mathbf{g}_p(\mathbf{x}) + \mathbf{w}_{o,p} \\ \mathbf{x}' &= \mathbf{f}(\mathbf{x}) + \mathbf{w}_x \end{aligned} \tag{1}$$

where \mathbf{x} and \mathbf{o}_p are respectively called hidden states and observations (the subscript p indicates the proprioceptive domain), and the letter w indicates noise terms sampled from Gaussian distributions. For simplicity, we considered just two temporal orders – although all the features we elucidate in the following hold for a system of generalized coordinates [58] – and we defined a likelihood function only for a single temporal order. We assume that the corresponding generative model is factorized as in Figure 1b, expressed in terms of *precisions* (or inverse variances) $\boldsymbol{\Pi}$. Note that we introduced a prior $\boldsymbol{\eta}_x$ over the hidden states, which is not generally used in continuous-time formulations, but it is the key element connecting different levels in discrete-time active inference [67] or PCNs [25] – as will be explained later. Also note that we used a generalized notation for instantaneous trajectories or paths, i.e., $\tilde{\mathbf{x}} = [\mathbf{x}, \mathbf{x}']$, where \mathbf{x} will be indicated in the following as the 0th order, and \mathbf{x}' as the 1st order. We highlighted in green and red respectively the input and output of the unit, namely the prior $\boldsymbol{\eta}_x$ and the observations \mathbf{o}_p .

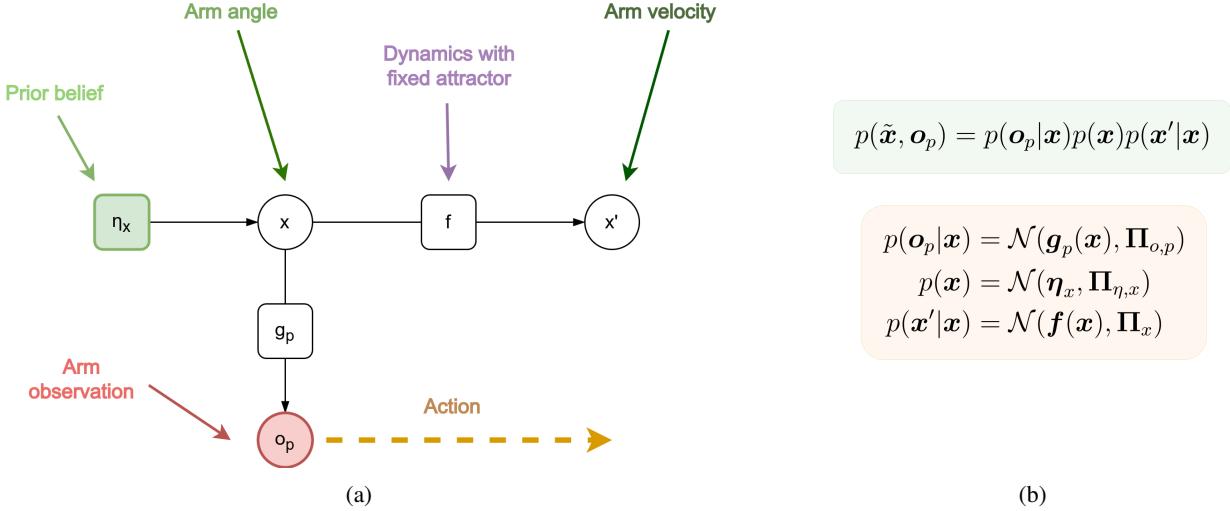


Figure 1: (a) Factor graph of a basic unit for static reaching. Variables and factors are indicated by circles and squares, respectively. Hidden states x (e.g., the arm angle) generate observations o_p (e.g., the arm proprioception) through the likelihood function g_p , and their 1st derivatives x' (e.g., the arm velocity) through a dynamics function f . In contrast to optimal control, here action follows observation prediction errors arising from a simple attractor ρ embedded in the model dynamics, or from a prior belief η_x over the arm angle. (b) Agent's generative model.

Exact computation of the posterior $p(\tilde{x}|o_p)$ is unfeasible since the evidence requires marginalizing over every possible outcome, i.e., $p(o_p) = \int p(\tilde{x}, o_p) d\tilde{x}$. For this reason, estimation of hidden states \tilde{x} is carried out through a variational approach [68], e.g., by minimizing the difference between a properly chosen approximate posterior $q(\tilde{x})$ and the true posterior. This difference is expressed in terms of a Kullback-Leibler (KL) divergence:

$$D_{KL}[q(\tilde{x})||p(\tilde{x}|o_p)] = \int_{\tilde{x}} q(\tilde{x}) \ln \frac{q(\tilde{x})}{p(\tilde{x}|o_p)} d\tilde{x} \quad (2)$$

The denominator $p(x|o_p)$ still depends on the marginal $p(o_p)$, but the KL divergence can be rewritten in terms of the log evidence and a quantity known as the *free energy* \mathcal{F} :

$$\mathcal{F} = \mathbb{E}_{q(\tilde{x})} \left[\ln \frac{q(\tilde{x})}{p(\tilde{x}|o_p)} \right] = \mathbb{E}_{q(\tilde{x})} \left[\ln \frac{q(\tilde{x})}{p(\tilde{x}|o_p)} \right] - \ln p(o_p) \quad (3)$$

Since the KL divergence is always nonnegative, the free energy provides an upper bound on surprise, i.e., $\mathcal{F} \geq \ln p(o_p)$. Hence, minimizing \mathcal{F} achieves the dual objective of keeping surprise low while estimating the true distribution. Assuming that the approximate posterior can be factorized into independent contributions, and further assuming that each contribution is Gaussian – i.e., $q(\tilde{x}) = \mathcal{N}(\tilde{\mu}_x, \tilde{P}_x)$, with generalized means $\tilde{\mu}_x$ and precisions \tilde{P}_x – the optimization process breaks down to the minimization of (precision-weighted) *prediction errors* in terms of the approximate posterior – see [35] for more details:

$$\begin{aligned} \varepsilon_{o,p} &= o_p - g_p(\mu_x) \\ \varepsilon_{\eta,x} &= \mu_x - \eta_x \\ \varepsilon_x &= \mu'_x - f(\mu_x) \end{aligned} \quad (4)$$

Then, the inference of the means $\tilde{\mu}_x = [\mu_x, \mu'_x]$ (also called *beliefs*) of the posterior over the hidden states is reduced to the following message passing:

$$\dot{\tilde{\mu}}_x = \begin{bmatrix} \dot{\tilde{\mu}}_x \\ \dot{\tilde{\mu}}'_x \end{bmatrix} = \mathcal{D}\tilde{\mu}_x - \partial_{\tilde{x}}\mathcal{F} = \begin{bmatrix} \mu'_x - \Pi_{\eta,x}\varepsilon_{\eta,x} + \partial_x g_p^T \Pi_{o,p} \varepsilon_{o,p} + \partial_x f^T \Pi_x \varepsilon_x \\ -\Pi_x \varepsilon_x \end{bmatrix} \quad (5)$$

where \mathcal{D} is an operator that shifts every derivative by one, i.e., $\mathcal{D}\tilde{\mu}_x = [\mu'_x, \mathbf{0}]$. This term arises because the generative model maintains a belief not over a static point, but over a dynamic trajectory, and only when the motion of the mean $\dot{\mu}_x$ equals the mean of the motion $\mathcal{D}\tilde{\mu}_x$, is the free energy minimized. In short, the inferential process does not involve matching a state (as in PCNs) but tracking a path [69]. Unpacking Equation 5, we note that the 0th order is subject to a forward error from the prior, a backward error from the likelihood, and a backward error from the dynamics function. On the other hand, the 1st order is only subject to the latter but in the form of a forward error. The belief is then updated via gradient descent, i.e., $\tilde{\mu}_{x,t+1} = \tilde{\mu}_{x,t} + \Delta_t \dot{\mu}_x$, where Δ_t is a time constant.

How can this agent perform a simple reaching movement? As highlighted in Figure 1a, we can encode the arm angle and velocity as generalized hidden states. We will talk later about the relation between proprioceptive and exteroceptive domains; for now, we consider a single DoF that has a univocal mapping between the joint angle of the arm and the Cartesian position of the hand. Indicating the target to reach by ρ , we can define the following dynamics function:

$$f(x) = \rho - x \quad (6)$$

expressing a simple attractor toward the target [53, 70, 71, 72, 73]. These dynamics do not exist in the actual generative process, and it is indeed this discrepancy that forces the environment to conform to the agent’s beliefs. Specifically, Equation 6 means that the agent thinks its hand will be pulled toward the target with a strength proportional to the precision Π_x . In fact, the attractor affects the belief update through the dynamics prediction error ε_x , expressing a difference between the estimated velocity μ'_x and the one predicted by the agent through the dynamics function f . Note that this error appears in both temporal orders: in brief, ε_x imposes a trajectory at the 1st order which in turn affects the 0th order directly through μ'_x , and indirectly through the gradient $\partial_x f$.

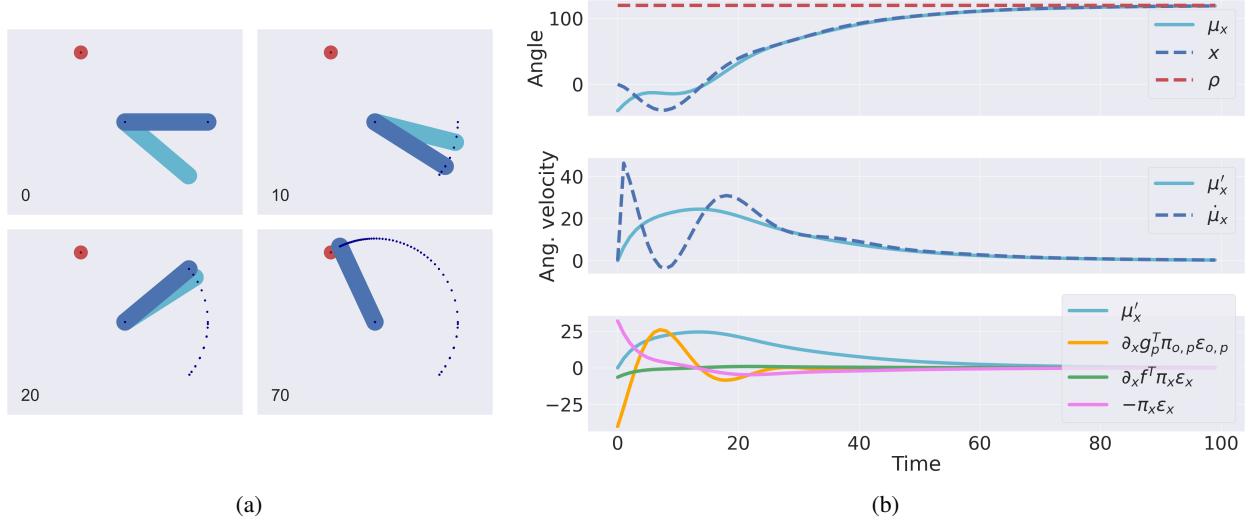


Figure 2: (a) In this task, the agent (a single DoF) has to reach a target angle represented by the red circle. Estimated and real arms are displayed in cyan and blue, respectively. Here, $\pi_{\eta,x} = 0$, $\rho = 120^\circ$, and μ_x was initialized to -40° . The time step is indicated in the bottom left corner of each frame. Since the belief was initialized at a negative value, the likelihood initially pulls the arm toward the wrong direction before adapting to the dynamics attractor. (b) The top graph shows the evolution of the real angle x , its belief μ_x , and the target angle ρ . The middle graph shows the evolution of the belief of the velocity μ'_x and the belief derivative $\dot{\mu}_x$. The bottom graph shows the evolution of all the components that comprise the belief update: the belief of the velocity μ'_x , the likelihood gradient $\partial_x g^T \Pi_{o,p} \varepsilon_{o,p}$, the dynamics gradient $\partial_x f^T \Pi_x \varepsilon_x$, and the weighted dynamics prediction error $-\Pi_x \varepsilon_x$. The latter has been plotted to compare its magnitude with the other components, although affecting the 1st temporal order.

The interactions between these quantities can be better understood from Figure 2, showing a reaching movement with the defined dynamics function and the trajectories of the agent’s generative model. Here, the belief is subject to two different forces: a likelihood gradient pushing it toward what it is currently perceiving (i.e., the real angle), and the other components that steer it toward the biased dynamics (i.e., the target angle ρ). Note how in the third plot, of the three components that comprise the belief update, the backward error $\partial_x f^T \Pi_x \varepsilon_x$ has the smallest amplitude. While the exact interactions arising from the dynamics prediction error have yet to be analyzed, in the following we assume

that goal-directed behavior is achieved through the forward error at the 1st order $-\Pi_x \varepsilon_x$. An alternative would be to directly control the backward error without maintaining a belief of the 1st order [70], which however requires taking a gradient into account and may be more challenging when defining appropriate attractors. Finally see how, in the middle plot, the agent tries at every instant to minimize the difference between μ'_x and μ_x , thus tracking the actual path of the hidden states.

But how does the agent move in practice? As mentioned in the Introduction, action is the other side of the coin of the free energy principle, through which the agent samples those observations conforming to its prior beliefs. In fact, in addition to the perceptual inference typical of predictive coding, active inference assumes that organisms minimize free energy also by interacting with the environment; this minimization breaks down to an even simpler update that only depends on (proprioceptive) prediction errors $\varepsilon_{o,p}$. Since these prediction errors are generated from the agent's belief, this means that whenever the latter is biased toward some preferred state, movement naturally follows. There is thus a delicate balance between perception – in which prediction errors climb up the hierarchy to bring the belief closer to the observations – and action – in which prediction errors are suppressed at a low level so that the observations are brought closer to their predictions. However, there is an open issue regarding how active inference should be practically realized in continuous time. A few studies demonstrated that using exteroceptive information directly for computing motor commands could result in smoother movements and resolution of visuo-proprioceptive conflicts [32, 44, 64], and in fact some robotic implementations effectively used this approach [70, 71]. However, evidence seems to indicate that motor commands are generated by suppression of proprioceptive signals only [38, 37], which is already in the intrinsic reference frame needed for movement and thus results in easier inverse dynamics. For this reason, in the following we assume that movements are realized by minimizing the free energy with respect to proprioceptive prediction errors:

$$\dot{\mathbf{a}} = -\partial_a \mathcal{F} = -\partial_a \mathbf{g}_p^T \Pi_{o,p} \varepsilon_{o,p} \quad (7)$$

where $\partial_a \mathbf{g}_p$ performs an inverse dynamics from proprioceptive predictions to motor commands \mathbf{a} , likely to be implemented by classical spinal reflex arcs. As a last note, the actions can also depend on multiple orders – velocity, acceleration, and so on – allowing more efficient movement and control [74, 75, 76, 77], but since it is beyond our scope, we only minimize the 0th order. Nonetheless, 1st-order movements – e.g., maintaining a constant velocity – are still possible by specifying appropriate dynamics of the hidden states.

2.2 Tracking objects

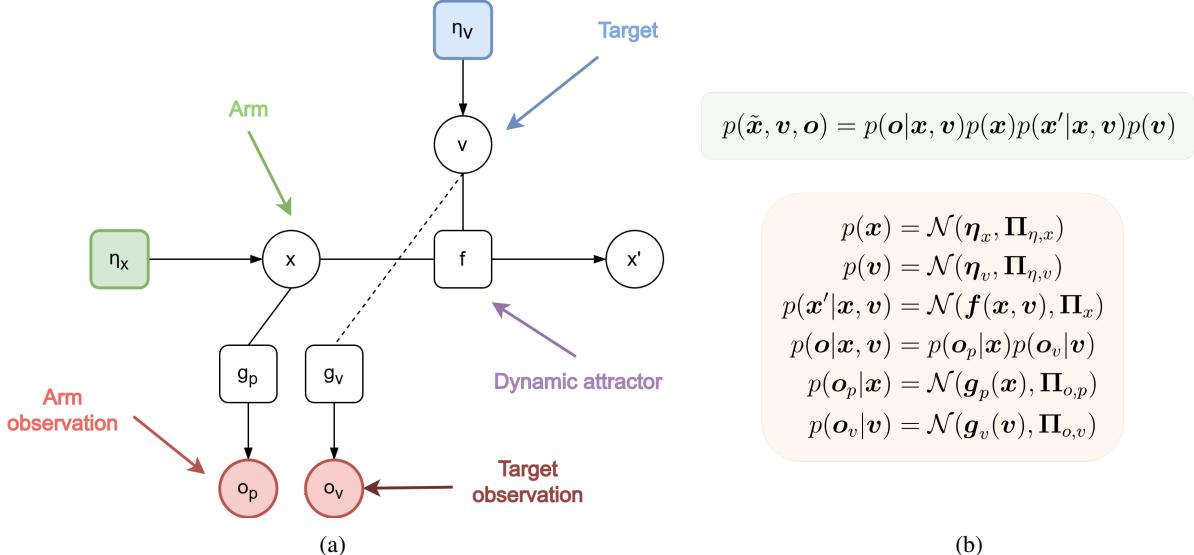


Figure 3: (a) The target is now encoded in the hidden causes v , generating a dynamic attractor for object tracking. In fact, both hidden states and hidden causes generate predictions through proprioceptive and visual likelihood functions g_p and g_v , and both concur in estimating the 1st-order hidden states x' . (b) Agent's generative model.

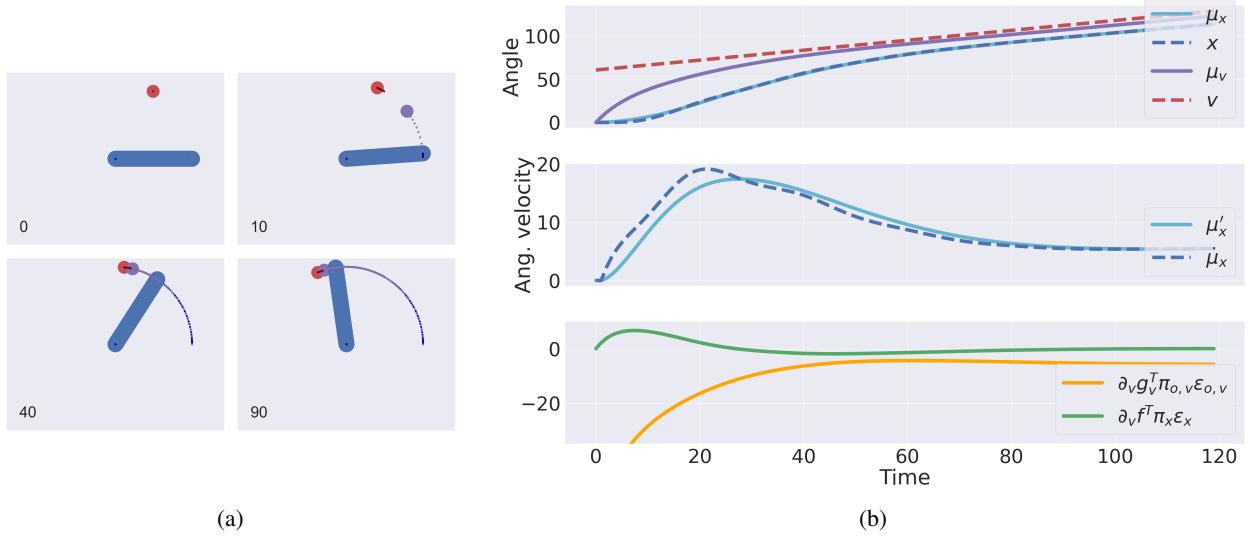


Figure 4: (a) In this task, the agent has to track a target angle rotating at a constant velocity. Estimated and real targets are displayed in purple and red, respectively. Here, $\Pi_{\eta,x} = 0$, $\Pi_{\eta,v} = 0$, v was initialized to 60° , and both μ_x and μ_v were initialized to 0° . Here, the belief of the hidden causes pulls the belief of the hidden states with it while approaching the real target angle. (b) The top graph shows the evolution of the real angle x , its belief μ_x , the target angle v , and its belief μ_v . The middle graph shows the evolution of the belief of the velocity μ'_x and the belief derivative $\dot{\mu}_x$, as before. The bottom graph shows the evolution of all the components that comprise the hidden causes update: the likelihood gradient $\partial_v g_v^T \Pi_{o,v} \epsilon_{o,v}$, and the dynamics gradient $\partial_v f^T \Pi_x \epsilon_x$. Note how, in the middle plot, the estimated 1st temporal order stabilizes to a non-zero value as the agent rotates with a constant angular velocity.

The simple agent defined in the previous section can only realize fixed trajectories embedded in the dynamics function, so how can it track moving objects? This is usually done by introducing a key concept in active inference, the *hidden causes* v , which link hierarchical levels and specify how the dynamics function evolves. In the active inference literature of motor control, they are also used to encode the target to be reached [32, 78, 63, 79], as depicted in Figure 3a. Considering the target as a causal variable for the hidden states and sensory observations makes sense from an active perspective whereby “it is an object I want to reach that generates my movements”. Now, the agent’s generative model becomes that of Figure 3b. Note that there are two priors, one over the hidden states and another over the hidden causes, respectively denoted by η_x and η_v . Also, both dynamics and likelihood functions depend on the hidden causes, and we assumed a further factorization for the likelihood, where o_p and o_v denote the arm and target observations, respectively. For simplicity, we assume that the visual likelihood function g_v is a simple identity generating the target angle. It is through the connection between hidden causes and observations that the agent can operate in dynamic environments. In fact, we can define the following dynamics function:

$$f(x, v) = v - x \quad (8)$$

where we just replaced the static target ρ with the hidden causes. As for the hidden states, we define a (Gaussian) approximate posterior over the hidden causes $q(v) = \mathcal{N}(\mu_v, P_v)$, with mean μ_v and precision P_v . Then, the mean is updated according to:

$$\dot{\mu}_v = -\partial_v \mathcal{F} = -\Pi_{\eta,v} \epsilon_{\eta,v} + \partial_v g_v^T \Pi_{o,v} \epsilon_{o,v} + \partial_v f^T \Pi_x \epsilon_x \quad (9)$$

where we defined the following observation and prior prediction errors in terms of the approximate posterior:

$$\begin{aligned} \epsilon_{o,v} &= o_v - g_v(\mu_v) \\ \epsilon_{\eta,v} &= \mu_v - \eta_v \end{aligned} \quad (10)$$

As evident, the hidden causes are subject to a prior prediction error, a backward dynamics error, and a backward likelihood error – similar to the update of the hidden states, except that this inference is over a state and not a path. Via the backward likelihood error, the agent can correctly estimate the target configuration whenever it moves, as shown in the tracking simulation of Figure 4. Concerning the dynamics prediction error, it can now flow into two different

pathways: specifically, the role of the gradients $\partial_x \mathbf{f}$ and $\partial_v \mathbf{f}$ are to respectively infer the state and the cause that may have generated a particular velocity; their actual role will be clear in Chapter 4.

2.3 Intention modulation and object affordances

Although capable of operating in dynamic contexts, the last approach still reproduces a simple scenario in which a target has no internal dynamics and always has the role of a cause for a hidden state. In other words, it does not permit modeling realistic tasks such as a pick-and-place operation, where an object is first the cause of a reaching and grasping movement, but then it becomes the consequence of another cause such as a goal position, resulting in a placing movement; critically, it does not allow to model a task wherein not only the dynamics of the self, but also the dynamics of the target must be learned (e.g., if a moving object should be grasped on the fly, the agent should infer its trajectory to anticipate where it will fall).

It follows that to operate in a complex environment, the agent must (i) maintain complete representations for each entity that it wants to interact with, and (ii) flexibly assign causes and consequences for the next movement depending on the current context – in a similar way to policies in discrete models, as will be explained later. Therefore, we first encode multiple environmental entities as *potential body configurations* in the hidden states, i.e., $\mathbf{x} = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N]$, where \mathbf{x}_0 is the actual body configuration (as before), and N is the number of entities [64]. Consequently, the factorized likelihood function generates a proprioceptive observation for the first component \mathbf{x}_0 , and visual observations for each entity:

$$\mathbf{o} = [\mathbf{o}_p, \mathbf{o}_{v,1}, \dots, \mathbf{o}_{v,N}] = [\mathbf{g}_p(\mathbf{x}_0), \mathbf{g}_v(\mathbf{x}_1), \dots, \mathbf{g}_v(\mathbf{x}_N)] \quad (11)$$

Here, visual observations are assumed to be in the Cartesian domain, so the visual likelihood function \mathbf{g}_v generates the hand positions of the potential configurations through forward kinematics. This structure is similar to the previous model, except that the target is now embedded in the hidden states along with the hand, and that there is no connection between hidden causes and observations. We could define a similar factorization for the hidden causes and dynamics function, such that each entity would have an independent dynamics biased by a specific cause (e.g., where the arm or the target will be in the future); however, this is of limited use in a pick-and-place operation that demands interaction between entities. We hence compute an intentional state with a single function, such as:

$$\mathbf{i}_m(\mathbf{x}) = \mathbf{W}_m \mathbf{x} + \mathbf{b}_m \quad (12)$$

The weights \mathbf{W}_m perform a linear transformation of the hidden states that combines every entity, while the bias \mathbf{b}_m imposes a static configuration [80]. Equation 12 can be realized through simple neural connections, wherein the weights are encoded as synaptic strengths and the bias represents the threshold needed to fire a spike, and both can be known a priori or learned via sensory evidence. An error is then computed between this intentional state and the current one:

$$\mathbf{e}_{i,m} = \mathbf{i}_m(\mathbf{x}) - \mathbf{x} \quad (13)$$

This vector has the same role as the attractor of Equation 8, but now it points toward a function of the hidden states. Finally, we define the following dynamics function:

$$\mathbf{f}_m(\mathbf{x}, v_m) = v_m \mathbf{e}_{i,m} \quad (14)$$

multiplying the error by a single-value hidden cause v_m . Thus, the latter is not intended as an explicit trajectory prior over the hidden states (e.g., encoding where my arm will be in the future), whose role is now delegated to the bias \mathbf{b}_m ; but as an attractor gain, whereby a high value implies a strong force toward the future state. Since \mathbf{i}_m is used to define a path for the hidden states aiming to produce a desired configuration, we call it an *intention*.

The dynamics function of Equation 14 is not composed of segregated pathways as for the likelihood, but affects all the environmental entities at once – e.g., it computes a trajectory for the arm depending on the target. The steps performed by the agent during a reaching movement are the following: (i) the 0th order imposes a trajectory to the 1st order and generates a sensory prediction; (ii) the 0th order infers the consequences of its predictions, hence it is now biased toward both the intentional state and the observation; (iii) a proprioceptive prediction is generated from this new biased position, eventually driving action. This approach can be seen as a generalization of [63] where, in the context of oculomotor behavior, the target and the center of gaze were encoded as hidden states, each with their own dynamics and attracted by a hidden location. Although limited compared to non-linear dynamics functions (e.g., obstacle avoidance can be realized via repulsive potentials [80]), with the specific form defined above – along with the hidden states factorization –

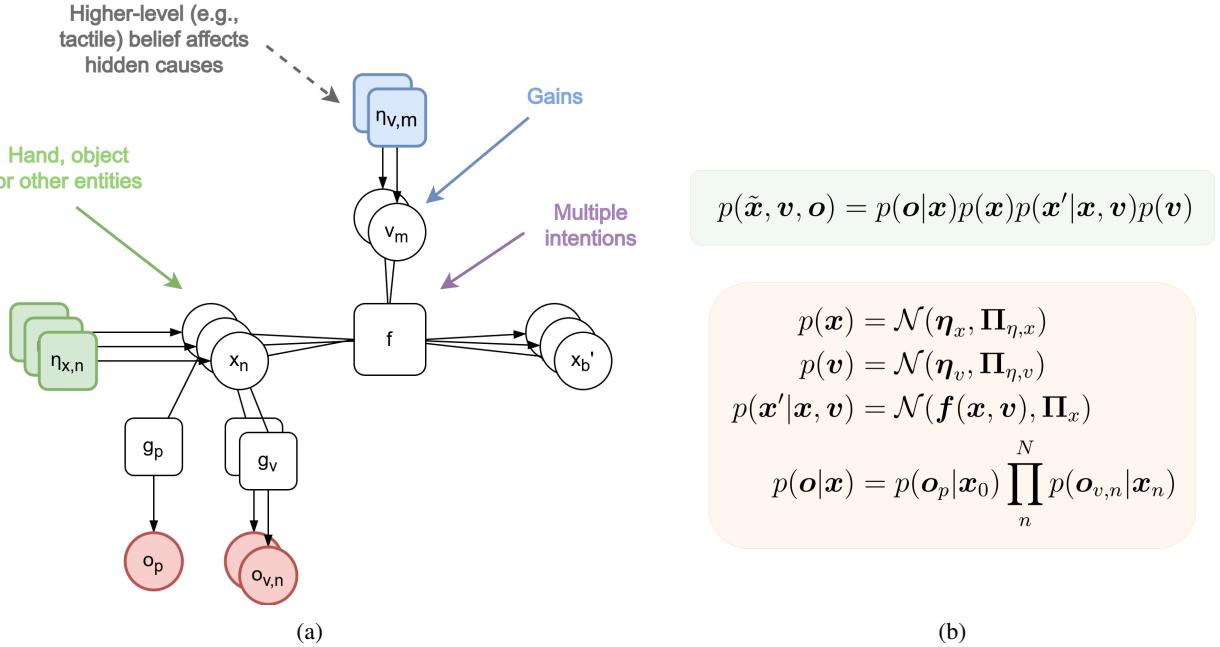


Figure 5: (a) Factor graph of the unit with object affordances. Hidden states are factorized into independent components that encode the actual bodily states and potential configurations related to the objects. The first component \mathbf{x}_0 generates proprioceptive predictions, while the successive components generate visual predictions of the objects. Every hidden cause v_m now defines an attractor gain expressing the strength of an agent’s intention (encoded as a distinct evolution of the world). These are combined to produce a trajectory η'_x , comprising all the body configurations \mathbf{x}_n . The transition between intentions can be achieved by a higher-level prior, e.g., a belief of tactile sensations. The weights \mathbf{W}_m of the intention can be used, e.g., to track moving objects, while the bias b_m realizes a static configuration. See [64, 66] for more details. (b) Agent’s generative model.

there is a high flexibility for complex interactions. Further, interpreting the hidden causes as a gain still makes sense from an active inference perspective because what is represented at a higher level is the intention to move at the target, while the target location is inferred at a lower level.

Taken alone, considering a hidden cause as an attractor gain may not seem so helpful. However, as depicted in Figure 5a, we can combine M intentions in the following way:

$$\eta'_x = \mathbf{f}(\mathbf{x}, \mathbf{v}) = \sum_m^M \mathbf{f}_m(\mathbf{x}, v_m) = \sum_m^M v_m(\mathbf{i}_m(\mathbf{x}) - \mathbf{x}) \quad (15)$$

In short, the agent entertains M distinct dynamic hypotheses of how the world may evolve. Trajectories $\mathbf{f}_m(\mathbf{x}, v_m)$ are separately computed from each intention i_m and with their respective gains; then, the final trajectory η'_x is found by combining all of them. The reason why we used the prior notation for the trajectory predictions will be clear in Chapter 4. Since each attractor $e_{i,m}$ is proportional to its hidden cause v_m , the latter lends itself to a parallelism with the policies of discrete models, as will be explained later: if v_m is set to 1 and all the others to 0, the hidden states will be subject only to intention m ; conversely, if multiple hidden causes are active, the hidden states will be pulled toward a combination of the corresponding intentions. This means that the hidden causes act both as attractor gains – expressing the absolute strength by which the belief is steered toward the desired direction – and as intention modulators – defining the relative strength between each trajectory. As a result, we have an additional modulation that combines with the dynamics precision $\boldsymbol{\Pi}_x$; their interactions will be explained in Chapter 4. The resulting behavior is similar to the Lotka-Volterra dynamics used in [81]. In the latter, the agent revisits in sequence points in attractor space linked to specific locations, and the defined dynamics ensure that only one attractor is active at any time. This permits modeling itinerant movements such as handwriting or walking.

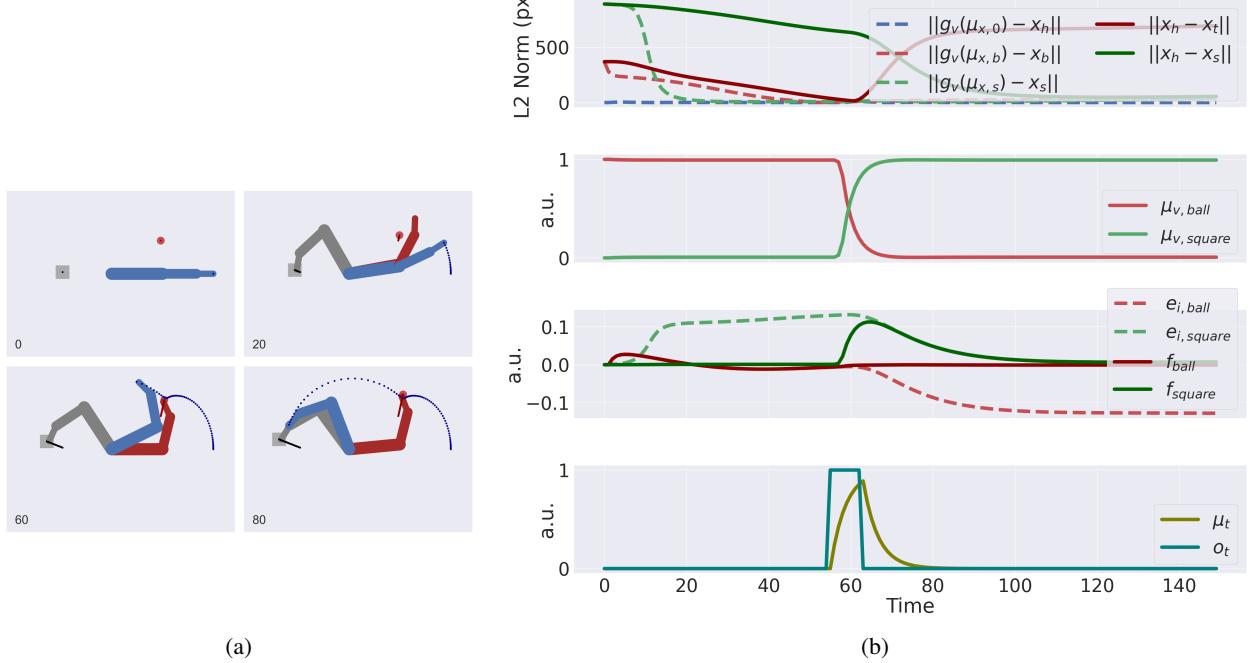


Figure 6: (a) In this task, the agent has to first touch the moving ball (in red) and then track the moving square (in grey). The transition is done by a tactile belief. Note how throughout the task, the agent maintains potential body configurations related to the two objects. See [64, 66] for more details. (b) The first graph shows the difference between the estimated and real hand positions (dotted blue), the difference between the potential hand position related to the ball and the actual ball position (dotted red), the difference between the potential hand position related to the square and the actual square position (dotted green), the difference between the real hand and ball positions (solid dark red), and between the real hand and square positions (solid dark green). The second graph shows the evolution of the hidden causes associated with the two intentions, $\mu_{v,ball}$ and $\mu_{v,square}$. The third graph shows the evolution of the attractors $e_{i,ball}$ and $e_{i,square}$, and the dynamics functions f_{ball} and f_{square} , regarding the first joint angle. The last graph shows the evolution of the tactile observation o_t , and its belief μ_t .

The generative model is shown in Figure 5b. The update rule for the 0th-order hidden states becomes:

$$\dot{\mu}_x = \mu'_x - \Pi_{\eta,x} \epsilon_{\eta,x} + \partial_x g_p^T \Pi_{o,p} \epsilon_{o,p} + \sum_n \partial_x g_v^T \Pi_{o,v,n} \epsilon_{o,v,n} + \partial_x f^T \Pi_x \epsilon_x \quad (16)$$

In particular, the dynamics prediction error:

$$\epsilon_x = \mu'_x - \eta'_x \quad (17)$$

realizes an average trajectory that the agent predicts for the current situation. This approach is effective for two reasons. First, it allows defining a composite movement in terms of simpler subgoals, which can be tackled separately; this can be helpful, for instance, if one has to analyze the behavior of an agent when subject to two or more opposing priors [64]. Second, a simple multi-step behavior without planning can be achieved [66], wherein one just needs to modulate the hidden causes. Transitions between continuous trajectories could then be realized by higher-level priors, e.g., a belief of tactile sensations. Third, and most importantly, it permits maintaining in parallel potential body configurations related to the objects to be manipulated – thus providing efficient transitions between movements – and encoding the objects according to their affordances and the agent’s intentions (e.g., grasping a cup by the handle or with the whole hand). These features are illustrated in the simulation of Figure 6, showing a two-step reaching task with moving objects.

3 Hierarchical models

So far, we have discussed several units with two kinds of inputs – a prior over the hidden states and a prior over the hidden causes – and one kind of output – the 0th-order observation. In this chapter, we focus on combining such units in

a single network to achieve a more advanced and efficient control. For this, we will make use of the first input, leaving the discussion about the second to the next chapter.

In hierarchical active inference models, units are arranged in layers so that the output of one layer provides input to the subordinate layers. This architecture permits representing complex data, such as convolutions models or nonlinear time series [57]. In motor control, a (deep) hierarchy of goals integrates control and motivational streams of the brain [82]. For robotics, a hierarchical kinematic model can be designed in continuous time, wherein each unit encodes a certain Degree of Freedom (DoF) in intrinsic and extrinsic reference frames [80]. This permits realizing advanced movements that involve simultaneous coordination of multiple limbs, e.g., moving with a glass in hand. This hierarchical structure can be generalized to perform homogeneous transformations between reference frames, e.g., perspective projections [83].

3.1 Intrinsic and extrinsic causes

The last unit presented affords a multi-step behavior in continuous time that accounts for object affordances and, to some extent, for dynamic elements of the environment. However, it can only estimate body configurations, while in real-life applications we constantly plan movements in the spatial domain. Further, it only generates visual predictions of an object related to a single DoF (e.g., the hand), while we generally deal with much more complex kinematic structures with different branches such as the human body. As in optimal control, continuous-time active inference considers three reference frames and two inversions: an *extrinsic* signal (e.g., encoding the Cartesian position of a target) is first transformed in an *intrinsic* signal (e.g., encoding the joint angles configuration corresponding to the hand at the target) through *inverse kinematics*, which is in turn converted to the actual motor control signals (e.g., joint torques) through *inverse dynamics* [84]. These two processes are also attributed to the human brain [85, 86, 87], but there is a substantial difference between optimal control and active inference regarding how they unfold in practice. As mentioned in the previous chapter, in active inference the motor commands are replaced by proprioceptive prediction errors that are suppressed through spinal reflex arcs [38]. As a consequence, inverse dynamics becomes easier because action is put aside and the agent has just to know the mapping from proprioceptive states to motor commands – see Equation 7.

But what about inverse kinematics? Recall the perspective that we mentioned in the previous chapter, i.e., that “it is an object I want to reach that generates my movements”. Turning optimal control upside down, active inference posits that action is driven by the proprioceptive consequences (e.g., changes in muscle lengths) of extrinsic causes (e.g., a limb position) [37]. Intuitively, one could model an extrinsic movement as in Figure 7a, i.e., with the following dynamics and likelihood functions:

$$\begin{aligned} f(\mathbf{x}, \mathbf{v}) &= \mathbf{J}^T(\mathbf{v} - \mathbf{T}(\mathbf{x})) \\ g_p(\mathbf{x}) &= \mathbf{x} \\ g_v(\mathbf{x}, \mathbf{v}) &= [\mathbf{T}(\mathbf{x}) \quad \mathbf{v}] \end{aligned} \tag{18}$$

where \mathbf{x} are the arm joint angles, \mathbf{v} is the target position to reach, \mathbf{T} is the forward kinematics returning the hand position, and \mathbf{J} is its Jacobian matrix.

The visual likelihood function g_v generates visual predictions for the hand and the target through forward kinematics and identity map, respectively. For goal-directed behavior, first an error between the target and hand positions is generated; then, an inverse kinematic model is embedded directly into the dynamics function, e.g., a Jacobian transpose or a pseudoinverse [70, 71, 72, 53, 32, 81, 79]. In other words, an extrinsic reference frame is inverted to generate an intrinsic state, which is in turn transformed again in the first domain to be compared with visual observations. As a result, forward and inverse kinematics are performed twice, once in the dynamics function and once in the forward and backward passes of visual inference, i.e., when backpropagating the visual prediction error $\epsilon_{o,v}$:

$$\partial g_v^T \epsilon_{o,v} = \begin{bmatrix} \mathbf{J} \\ \mathbf{1} \end{bmatrix} (\mathbf{o}_v - [\mathbf{T}(\mathbf{x}) \quad \mathbf{v}]) \tag{19}$$

If the predictions are not temporarily stored, this requires increased computational demand and memory. Additionally, there is an issue regarding biological plausibility: using sensory-level attractors within the dynamics function means that a unit is aware of part of the likelihood prediction – generally assumed to go all down to the sensorium – and its inverse mapping, which are lower-level features. Finally, with the model in Figure 7a the agent cannot easily express

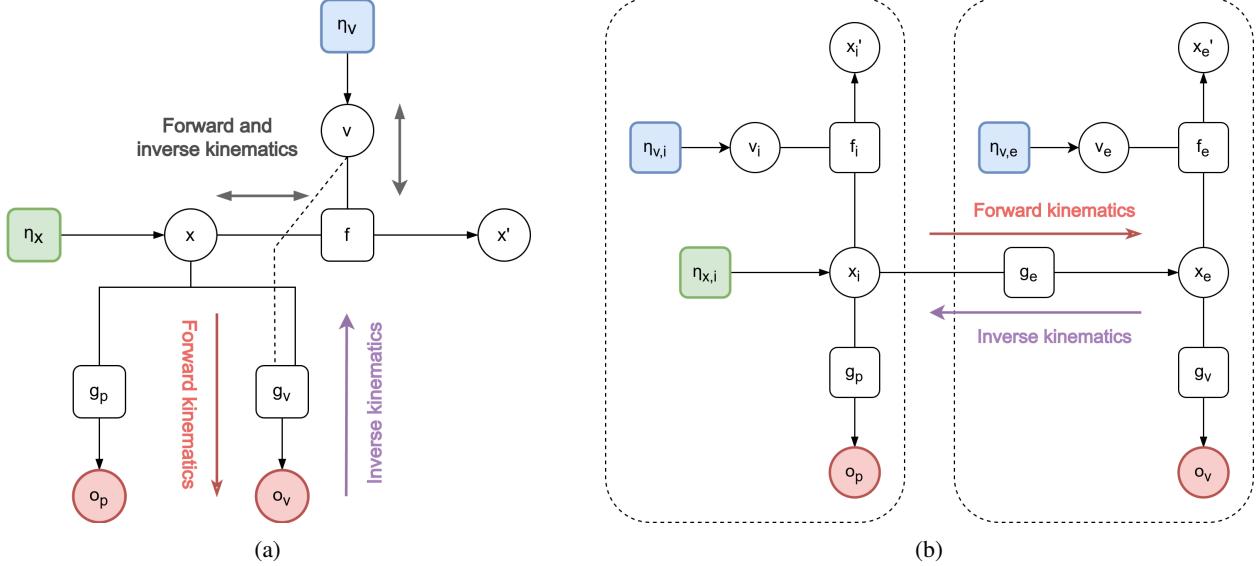


Figure 7: (a) Factor graph of an active inference model commonly used for kinematics. Hidden causes of a single level represent a target to reach, while hidden states define the joint angles of the kinematic chain, generating proprioceptive predictions through g_p and visual predictions through g_v . Forward and inverse kinematics are duplicated, further requiring the likelihood g_v to be embedded into the dynamics function. (b) Factor graph of an alternative hierarchical model for kinematics. Two different units (left and right blocks) encode information about arm joint angles and target position, respectively generating proprioceptive and visual predictions. The two levels are linked by the likelihood function g_e performing forward kinematics. Inverse kinematics and goal-directed behavior arise naturally through inference; moreover, both levels can express their own dynamics, affording more advanced control. (c) Generative model of (b).

paths in extrinsic coordinates needed, e.g., for realizing linear or circular motions, or for imposing constraints in both intrinsic and extrinsic domains such as when walking with a glass in hand.

We can instead exploit forward and inverse kinematics of Equation 19 and follow the natural flow of the generative process to avoid duplicated computations, as displayed in Figure 7b. The alternative generative model is displayed in Figure 7c. This model relies on two hierarchical levels, where an intrinsic unit (encoding the arm joint angles) is placed at the top and generates predictions through forward kinematics for an extrinsic unit (encoding the Cartesian position of the target) [80]:

$$\mathbf{x}_e = \mathbf{g}_e(\mathbf{x}_i) + \mathbf{w}_{o,e} = \mathbf{T}(\mathbf{x}_i) + \mathbf{w}_{o,e} \quad (20)$$

while the visual likelihood becomes a simple identity map, i.e., $\mathbf{g}_v(\mathbf{x}_e) = \mathbf{x}_e$. The goal-directed behavior from the dynamics function of Equation 18 arises naturally via backpropagation of the *extrinsic prediction error* $\epsilon_{o,e}$:

$$\partial \mathbf{g}_e^T \epsilon_{o,e} = \mathbf{J}^T (\boldsymbol{\mu}_{x,e} - \mathbf{T}(\boldsymbol{\mu}_{x,i})) \quad (21)$$

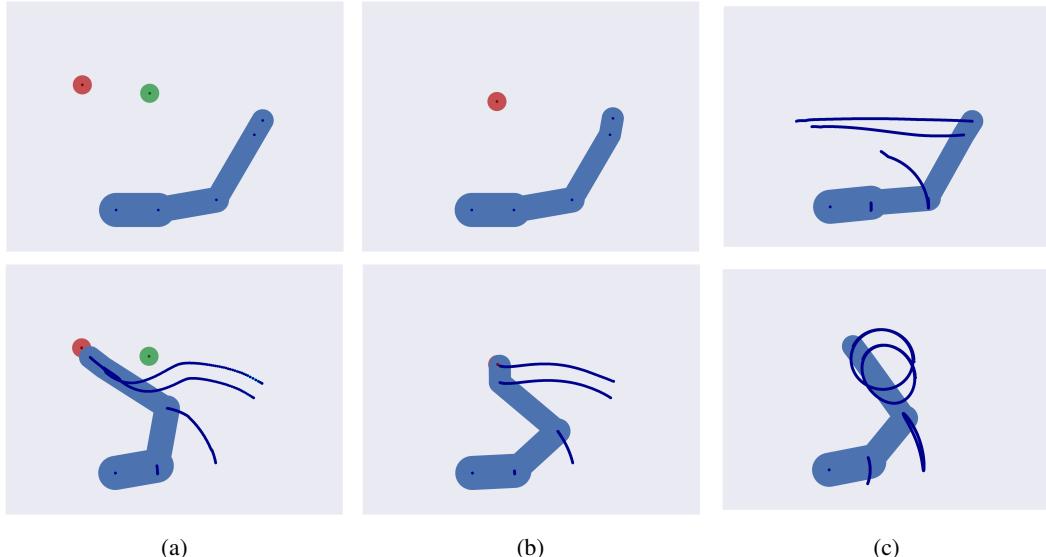


Figure 8: **(a)** In this task, the agent (a 4-DoF arm) has to reach the red target while avoiding the green obstacle; this is possible by specifying two (attractive and repulsive) functions at the extrinsic level. **(b)** In this task, the agent has to reach the red target while maintaining the same hand orientation (such as when walking with a glass in hand); this is possible by combining intrinsic and extrinsic constraints. **(c)** In this task, the agent has to perform linear (top) and circular (bottom) motions, possible by defining attractors at the 1st temporal order of the extrinsic hidden states. See [80] for more details.

Having a complete unit that deals with extrinsic information – which is thus not embedded into the hidden causes of the intrinsic unit – allows the agent to specify its dynamics, leading to an efficient decomposition between intrinsic and extrinsic attractors – i.e., $f_i(x_i, v_i)$ and $f_e(x_e, v_e)$ – and between proprioceptive and visual observations – as exemplified in the simulations of Figure 8. Note the similarity of Equation 21 with Equations 18 and 19: if in the model of Figure 7a we had two different forward and inverse kinematics either for goal-directed behavior or for predicting current observations, in the new model of Figure 7b what is compared with the observations already contains a bias toward intentional states, without the need for sensory-level attractors within the dynamics function. The extrinsic prediction error of Equation 21 becomes zero only when the extrinsic hidden states x_e match the predictions of the intrinsic unit; if such predictions are not met, $\varepsilon_{o,e}$ will flow through the hierarchy eventually generating an action. The update rules for the 0th orders of the intrinsic and extrinsic hidden states are the following:

$$\dot{\mu}_{x,i} = \mu'_{x,i} - \Pi_{\eta,x,i} \varepsilon_{\eta,x,i} + \partial_{x_i} g_p^T \Pi_{o,p} \varepsilon_{o,p} + \partial_{x_i} g_e^T \Pi_{o,e} \varepsilon_{o,e} + \partial_{x_i} f_i^T \Pi_{x,i} \varepsilon_{x,i} \quad (22)$$

$$\dot{\mu}_{x,e} = \mu'_{x,e} - \Pi_{o,e} \varepsilon_{o,e} + \partial_{x_e} g_v^T \Pi_{o,v} \varepsilon_{o,v} + \partial_{x_e} f_e^T \Pi_{x,e} \varepsilon_{x,e} \quad (23)$$

Although the generative model follows the forward flow of optimal control, the relationship between proprioceptive consequences and extrinsic causes peculiar to active inference still holds because the kinematic inversion regards a high-level process that manipulates abstract (intrinsic or extrinsic) representations, and both of them concur to generate low-level proprioceptive states. As Adams and colleagues note, “*The key distinction is not about mapping from desired states in an extrinsic (kinematic) frame to an intrinsic (dynamic) frame of reference, but the mapping from desired states (in either frame) to motor commands*” [38]. Having said this, there is a significant difference between the two models represented in Figure 7, which can be compared to the two supervised learning modes of predictive coding [25]: a forward mode that fixes the latent states to the labels and the observations to the data can generate highly accurate images of digits, while the inverse classification task is more difficult as there is no univocal mapping between labels and data; instead, a backward mode that fixes the latent states to the data and the observations to the labels achieves high performances on classification but falls short when generating images. Based on this, we can interpret the model of Figure 7a as a backward mode that would rapidly generate a proper kinematic configuration with the hand at the target, but that would hardly infer from proprioception the hand position needed to plan movements. Conversely, we can interpret the model of Figure 7b as a forward mode that would generate with high accuracy the hand position, but that would find it difficult to infer the kinematic configuration needed to actually realize movement.

3.2 A module for iterative transformations

The model in Figure 7b introduced a hierarchical dependency between two (intrinsic and extrinsic) levels, made possible by connecting hidden states. Instead, the typical approach in continuous-time active inference involves connections between hidden states and causes of a level and hidden causes (and not hidden states) of the subordinate level, as shown in Figure 9a. While this allows one to impose a trajectory for the unit below, specifying fixed setpoints to the 0th-order hidden states is not as straightforward, since the dynamics prediction error generated from the hidden causes would have to travel back to the previous temporal orders. As clear from Figure 7b, a connection between hidden states is of high utility when designing hierarchical models. In fact – as represented in Figure 9b – it is fundamental in defining the initial states of different temporal scales in discrete models, e.g., for pictographic reading [88] or linguistic communication [88]. Similar connections are used in standard PCNs, wherein each neuron of a level computes a combination of neurons of the level above passed to an activation function [25] – as shown in Figure 9c.

Recall that, in the previous kinematic model, the prediction $\mathbf{g}_e(\mathbf{x}_i)$ of the intrinsic hidden states acted as a prior for the extrinsic hidden states, while the latter acted as an observation for the intrinsic hidden states. Following this example, we use the observation of a level to bias the 0th-order hidden states of the level below directly:

$$\boldsymbol{\eta}_x^{(i+1)} \equiv \mathbf{g}^{(i)}(\boldsymbol{\mu}_x^{(i)}) \quad (24)$$

$$\mathbf{o}^{(i)} \equiv \mathbf{x}^{(i+1)} \quad (25)$$

As a result, the observation prediction error ε_o and the prior prediction error $\varepsilon_{\eta,x}$ of Equation 4 is expressed by the same variable:

$$\varepsilon_o^{(i)} = \varepsilon_{\eta,x}^{(i+1)} = \boldsymbol{\mu}_x^{(i+1)} - \mathbf{g}^{(i)}(\boldsymbol{\mu}_x^{(i)}) \quad (26)$$

where the hierarchical level is indicated with a superscript and lower levels are denoted by increasing numbers. We can then design a multiple-input and multiple-output system wherein a level imposes and receives priors and observations to several independent units, as in Figure 9d. The computation of the free energy in Equation 3 remains unchanged, and the update of the hidden states turns into the following:

$$\dot{\boldsymbol{\mu}}_x^{(i,j)} = \begin{bmatrix} \boldsymbol{\mu}_x'^{(i,j)} - \boldsymbol{\Pi}_o^{(i-1)} \varepsilon_o^{(i-1)} + \sum_l \partial_{x^{(i,j)}} \mathbf{g}^{(i,l)T} \boldsymbol{\Pi}_o^{(i,l)} \varepsilon_o^{(i,l)} + \partial_{x^{(i,j)}} \mathbf{f}^{(i,j)T} \boldsymbol{\Pi}_x^{(i,j)} \varepsilon_x^{(i,j)} \\ - \boldsymbol{\Pi}_x^{(i,j)} \varepsilon_x^{(i,j)} \end{bmatrix} \quad (27)$$

where the superscript notation (i,j) indicates the i th hierarchical level and the j th element within the same level. As evident, this is a similar connectivity to hierarchical dynamical models in temporal predictive coding [21, 89]. Here, the forward prediction error $\varepsilon_o^{(i-1)}$ combines the predictions from the units above and acts as prior (the gradient is absent since they are expressed in the same domain of the unit considered), while the backward prediction errors $\varepsilon_o^{(i,l)}$ contain the observations from the units below.

What advantages do deep hierarchical models carry compared to a shallow agent? Although the structure of Figure 7b affords a more advanced control with respect to the model in Figure 7a, its uses are still limited to solving simple tasks, e.g., performing operations with the hand. While simultaneous coordination of multiple limbs is possible, it would require complex dynamics functions, with complexity increasing with the number of joints and ramifications of the kinematic chain. Critically, a shallow agent would not be capable of capturing the hierarchical causal relationships inherent to the generative process, allowing one to predict and anticipate the local exchange of forces that would unfold during movement. As mentioned in the Introduction, a deep model is also required if one has to use tools for manipulation tasks. Besides roto-translations in forward kinematics, iterative transformations are also essential in computer vision – where an image can be subject to scaling, shearing, or projection – and, more in general, whenever changing the basis of a coordinate vector.

For these reasons, we can generalize the last model and construct an *Intrinsic-Extrinsic* (or IE) module [80, 83, 90]. This module is composed of two units and its role is to perform iterative transformations between reference frames. In brief, a unit $\mathcal{U}_e^{(i-1)}$ encodes a signal in an extrinsic reference frame, while a second unit $\mathcal{U}_i^{(i)}$ contains a generic intrinsic transformation. Applying the latter to the first signal returns a new extrinsic reference frame embedded in a unit $\mathcal{U}_e^{(i)}$.

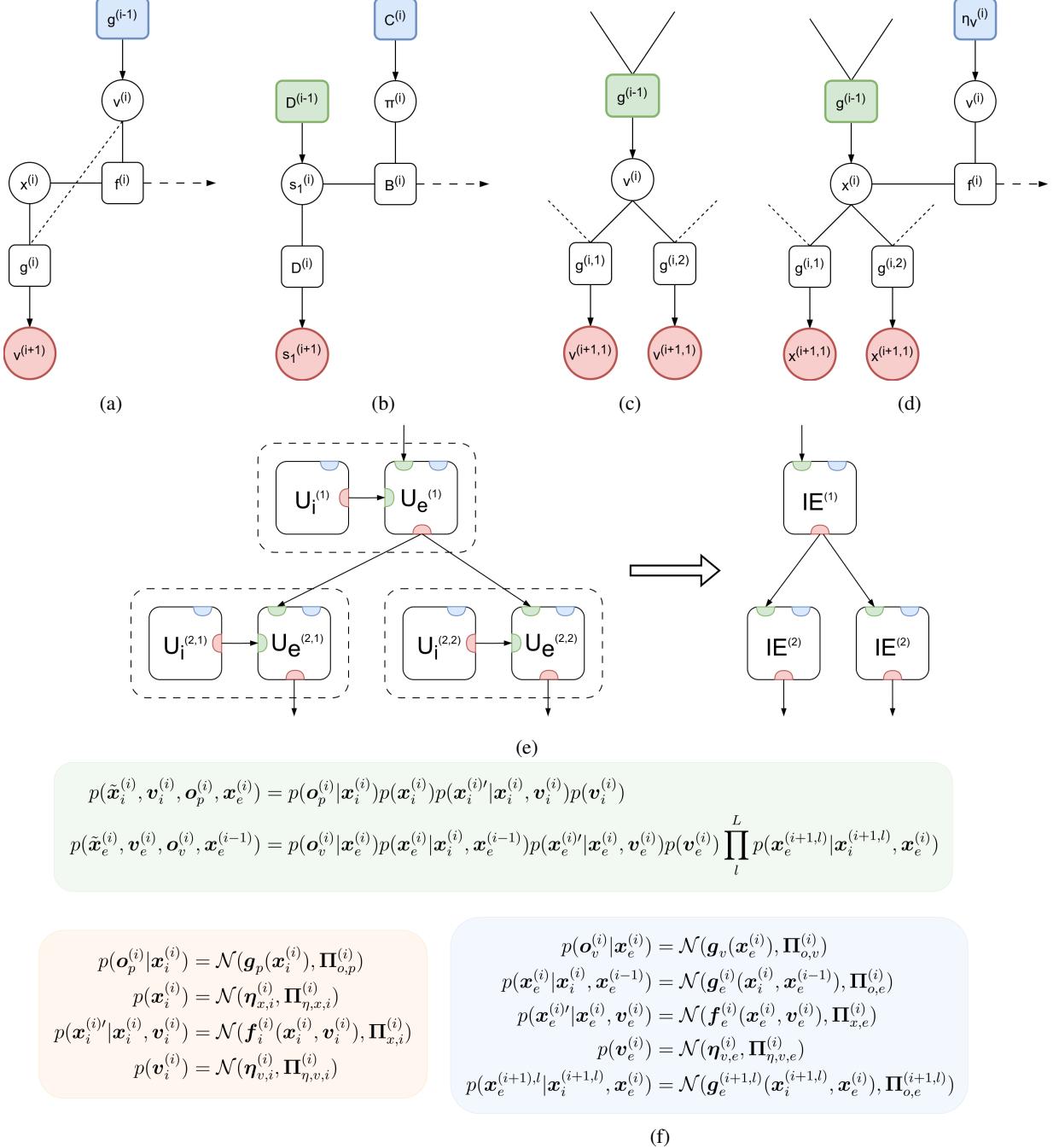


Figure 9: (a) Factor graph of the typical connection between two continuous levels. Hidden states $x^{(i)}$ and hidden causes $v^{(i)}$ of level i generate – through the likelihood function $g^{(i)}$ – the hidden causes $v^{(i+1)}$ of level $i + 1$. (b) Connections between two discrete levels. Hidden states $s_1^{(i)}$ of level i generate – through the prior matrix $D^{(i)}$ – the hidden states $s_1^{(i+1)}$ of level $i + 1$. We will cover discrete models in the next chapter. (c) Connections between two levels in PCNs. The likelihood function $g^{(i)}$ performs a simple combination of neurons, passed to a nonlinear activation function. Standard PCNs only permit representing the causal structure of a system, without modeling internal dynamics. (d) Factor graph of a level with multiple inputs and outputs from independent units, with similar connectivity to temporal predictive coding. The observation $\mathbf{o}^{(i,j)}$ becomes the 0th-order hidden state $x^{(i+1,j)}$ of the level below, while the prior $\eta_x^{(i,j)}$ becomes the 0th-order hidden state $x^{(i-1,j)}$ of the level above. (e) A network of IE modules. An extrinsic $\mathbf{x}_e^{(i-1)}$, along with an intrinsic signal $\mathbf{x}_i^{(i,j)}$ (e.g., angle for rotation or length for translation), is passed to a function $g_e^{(i,j)}$, generating a new extrinsic signal $\mathbf{x}_e^{(i,j)}$. (f) Generative model of (e).

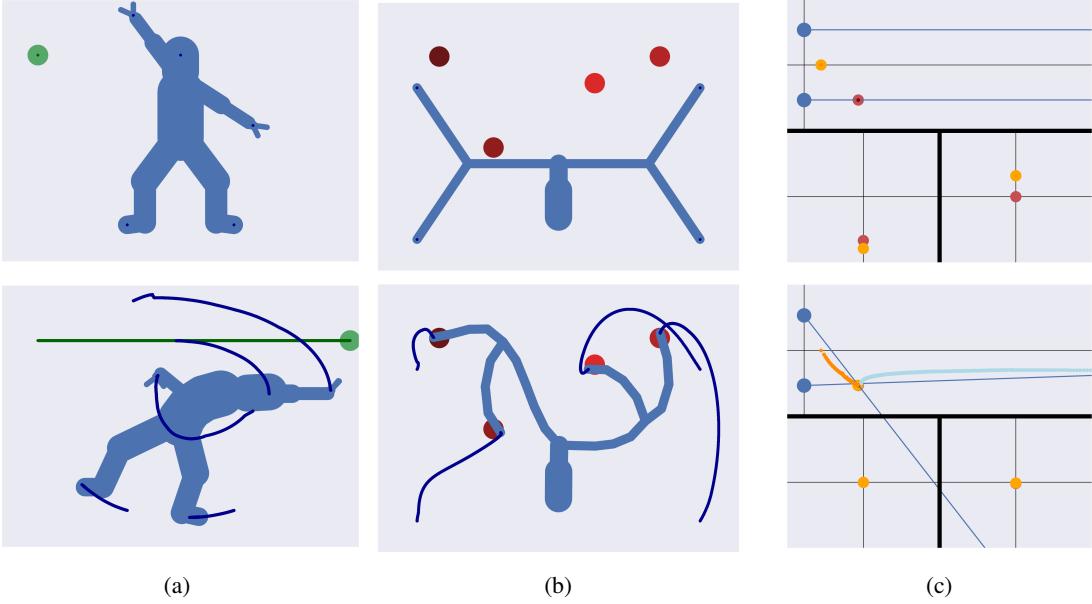


Figure 10: **(a)** In this task, the agent (a 23-DoF human body) has to avoid a moving obstacle; this is possible by defining a repulsive attractor for each extrinsic level. **(b)** In this task, the agent (a 28-DoF kinematic tree) has to reach four target locations with the extremities of its branches. See [80] for more details. **(c)** In this task, the agent (the two blue eyes on the left of the top view) has to infer the depth of the red object while fixating on it. The inferred position (and its trajectory) is represented in orange, while the blue trajectory is the center of fixation of the eyes. The bottom two frames show the object projection in the eye planes. See [83] for more details.

Then, we can define a likelihood function \mathbf{g}_e such that:

$$\mathbf{x}_e^{(i)} = \mathbf{g}_e^{(i)}(\mathbf{x}_i^{(i)}, \mathbf{x}_e^{(i-1)}) + \mathbf{w}_{o,e}^{(i)} = \mathbf{T}^{(i)}(\mathbf{x}_i^{(i)})^T \mathbf{x}_e^{(i-1)} + \mathbf{w}_{o,e}^{(i)} \quad (28)$$

where $\mathbf{w}_{o,e}^{(i)}$ is a noise term, while $\mathbf{T}^{(i)}$ here indicates a linear transformation matrix, which can express forward kinematics as in the previous examples or non-affine transformations. Backpropagating the extrinsic prediction error $\varepsilon_{o,e}^{(i)} = \boldsymbol{\mu}_{x,e}^{(i)} - \mathbf{g}_e(\boldsymbol{\mu}_{x,i}^{(i)}, \boldsymbol{\mu}_{x,e}^{(i-1)})$ leads to simple belief updates:

$$\begin{aligned} \partial_{\mathbf{x}_e^{(i-1)}} \mathbf{g}_e^{(i)T} \varepsilon_{o,e}^{(i)} &= \mathbf{T}^{(i)T} \varepsilon_{o,e}^{(i)} \\ \partial_{\mathbf{x}_i^{(i)}} \mathbf{g}_e^{(i)T} \varepsilon_{o,e}^{(i)} &= \partial_{\mathbf{x}_i^{(i)}} \mathbf{T}^{(i)} \odot [\varepsilon_{o,e}^{(i)} \boldsymbol{\mu}_{x,e}^{(i-1)T}] \end{aligned} \quad (29)$$

where \odot is the element-wise product. These equations express the most likely intrinsic and extrinsic states that may have generated the new reference frame. As shown in Figure 9e, modules are linked through the extrinsic units $\mathcal{U}_e^{(i)}$, while $\mathcal{U}_i^{(i)}$ performs an internal operation and does not contribute to the hierarchical connectivity. For motor control, we can realize a hierarchical multiple-output system, wherein the intrinsic hidden states $\mathbf{x}_i^{(i,j)}$ of a level encode a pair of joint angle and limb length of a single DoF. Iteratively applying roto-translations to an origin (e.g., body-centered) reference frame $\mathbf{x}_e^{(0)}$ – consisting of a Cartesian position and an absolute orientation – will determine the kinematic configuration of the agent in terms of extrinsic coordinates [80]. The generative model for the intrinsic and extrinsic units of a single IE module at level i are displayed in Figure 9f. Compared to the single-level generative model of Figure 7c comprising every joint angle of the agent’s body, here we have a dependency between levels encoding distinct joint angles. In particular, an additional term that links the extrinsic unit with the previous level, i.e., $p(\mathbf{x}_e^{(i)} | \mathbf{x}_i^{(i)}, \mathbf{x}_e^{(i-1)})$, and multiple distributions $p(\mathbf{x}_e^{(i+1,l)} | \mathbf{x}_i^{(i+1,l)}, \mathbf{x}_e^{(i)})$ generating extrinsic predictions for different modules at the next level.

At this point, we can easily express how every single joint and limb would evolve, affording a highly advanced control as demonstrated by the simulations of Figures 10a and 10b. Besides modeling limb dynamics, the IE modules can also be applied to other linear transformations, e.g., perspective projections. As displayed in Figure 10c, this can be useful for estimating the depth of an object via parallel predictions (e.g., from the eyes or multiple cameras) [83] – a process that active inference casts in terms of target fixation and hypothesis testing [78]. The modularity of this architecture allows the agent to define dynamic attractors in the 2D projected planes, in the 3D reference frames of the eyes, or as simple vergence-accommodation angles. This approach also has some analogies with Active Predictive Coding [91] and Recursive Neural Programs [92], which addressed the part-whole hierarchy learning problem in computer vision by recursively applying reference frame transformations to parts of a scene.

3.3 The self, the objects, and the others

Describing Figure 7b, we passed over a critical mechanism introduced at the beginning: the characterization of object affordances. Recall that the hidden states encoded in parallel not only the self but also other environmental entities; however, the agent’s model can now express the generative process hierarchically. This is described by the following likelihood function:

$$g_e^{(i)}(\mathbf{x}_i^{(i)}, \mathbf{x}_e^{(i-1)}) = \begin{bmatrix} \mathbf{T}^{(i)}(\mathbf{x}_{i,0}^{(i)})^T \mathbf{x}_{e,0}^{(i-1)} & \mathbf{T}^{(i)}(\mathbf{x}_{i,1}^{(i)})^T \mathbf{x}_{e,1}^{(i-1)} & \dots & \mathbf{T}^{(i)}(\mathbf{x}_{i,N}^{(i)})^T \mathbf{x}_{e,N}^{(i-1)} \end{bmatrix} \quad (30)$$

in which every IE module of the hierarchy has distinct factors for the self and every entity. For the self, this has a simple explanation, i.e., it just generates, one after the other, the positions of every segment of the kinematic chain depending on its joint angles. Concerning an object, we could encode its Cartesian position by attaching its visual observation to a second factor of the hidden states at a specific level. If the generative model has the same hierarchy for both the self and the object, backpropagating the extrinsic prediction errors of this second component will eventually infer a potential agent’s configuration in relation to the object, as before. For instance, if the object is linked to the last (i.e., hand) level, this would represent the hand at the object location, while all the previous levels would represent appropriate intermediate positions and angles generating that final location. In other words, the additional factorizations of hidden states and likelihoods reflect here a (deep) hierarchical configuration of the self that the agent thinks to be suitable for object manipulation. Since each level can express some dynamics through its hidden causes, the inference of this potential configuration is steered to match the object’s affordances and the agent’s intentions. As will be shown in the next chapter, this permits flexible adaptation of the kinematic chain depending on the circumstances, as well as representing the hierarchical structures of objects (e.g., tools). The inferred beliefs would be subject only to exteroceptive information from the objects, while proprioceptive states would be used only to update the agent’s belief of its current configuration.

Besides modeling object affordances, this strategy is also useful in multi-agent contexts. One could maintain a hierarchical generative model regarding the kinematic chain of another agent, which would be inferred by exteroceptive observations about all its positions and joint angles, starting from a different body-centered reference frame. As shown in Figure 11, the goal-directed method used for external objects reflects in this case as well: the agent could represent, by a parallel hierarchical pathway, a second agent in relation to itself, expressing a particular kind of interaction (e.g., the hand of the second agent in terms of its own, resulting in a shaking action). These two cases could be interpreted, from a biological perspective, as simulating the functioning of mirror neurons, firing whenever a subject executes a voluntary goal-directed action or when that action is performed by other subjects [93]. Building an internal model with the kinematic chain of the others – both per se and in relation to the self – could be critical to predict (thus, to understand) their intentions. In this view, neural activity results because the agent makes constant predictions over their kinematic structures depending on its hypotheses and the current context [94, 81].

The relationships between the self, the objects, and other agents under active inference may be better understood from the simulation of Figure 12, showing two agents with incompatible goals that depend on each other. Here, both agents are able to infer parallel representations of different kinematic chains, using an effective decomposition of potential and real configurations. Note how one’s current belief is always in between the intentional state to be realized and the actual configuration; this speaks of one of the fundamental aspects of active inference, i.e., that our beliefs never really reflect the state of the affairs of the world, but are always biased toward preferred states – eventually driving action. In general, bodily states, objects, or other agents can all be manipulated in reference frames appropriate for a specific context; this

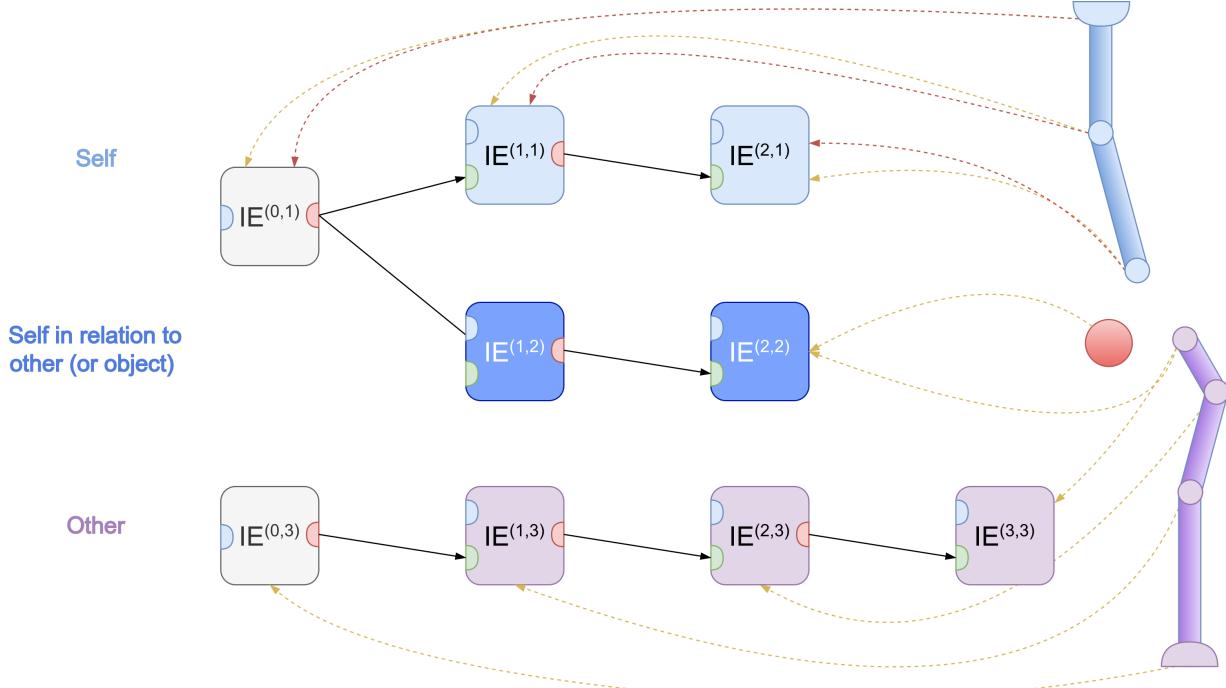


Figure 11: Interactions between an agent (a 2-DoF light-blue arm), another agent (a 3-DoF purple arm), and an object (a red ball). Generative model of the first agent, composed of three parallel pathways representing the kinematic structures of both agents. For clarity, lateral connections among the model components of each level are not shown. Both the self and the other (or object) in relation to the self depend on the same body-centered reference frame. The first component is subject to both proprioception (yellow dotted lines) and exteroception (red dotted lines), while the other two components are inferred via exteroception only. In this case, the interaction with the second agent just depends on the observation of its last level, leading, e.g., to a hand-shaking action.

is in line with the hypothesis that cortical columns use object-centered reference frames to encode external elements and more abstract entities [95].

4 The hybrid unit

The continuous-time hierarchical models presented so far lack effective usability in the real world: although they can represent any future trajectory - which implies a degree of planning - they do not possess an explicit model of future states nor can they choose among alternative trajectories. In this chapter, we turn to the problem of how to integrate discrete decision-making into continuous motor control. In doing this, we revisit the basic unit of the first chapter, finally using the second input – the prior over the hidden causes.

Active inference in discrete state-space [96, 97] – generally attributed to the cerebral cortex, especially prefrontal areas [98], along with corticostriatal loops – exploits the structure of Partially Observable Markov Decision Processes (POMDPs) to plan abstract actions over expected sensations. This (active) inference relies on the minimization of the *expected free energy*, i.e., the free energy that the agent expects to perceive in the future. The expected free energy can be unpacked into two terms resembling the two classical aspects of control theory, exploration and exploitation – which here naturally arise; these respectively correspond to an uncertainty-reducing term, and a goal-seeking term that finds a sequence of actions toward the agent’s prior belief.

Further, so-called *mixed* or *hybrid* models [49, 67] combine the potentialities of a discrete model with the inference of continuous signals, allowing robust decision-making in changing environments. While the theory of *Bayesian model reduction* [99, 100, 101, 102] provides efficient communication between the two models, this unified approach has not enjoyed many practical implementations for the time being [35, 49, 67, 103, 104, 105, 78]. An open issue regards how to deal with highly dynamic environments: hybrid models usually perform a comparison between static priors,

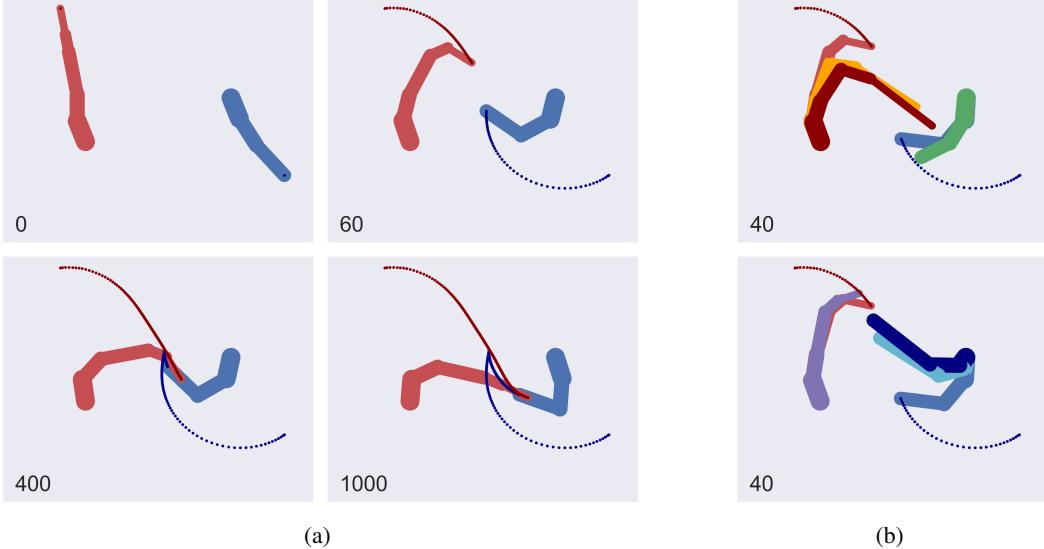


Figure 12: (a) In this task, two agents with different kinematic chains (respectively of 5 and 3 DoF) have two incompatible goals: the first agent (in red) has to reach the elbow of the second agent (in blue), while the second agent has to reach the hand of the first agent. Note that after an initial approaching phase from both, the second agent gradually retracts trying to touch the hand of the first agent. (b) Beliefs of both agents, compared with their actual configurations (displayed in light red for the first agent, and in light blue for the second). The top and bottom graphs respectively show the beliefs of the first and second agents. Specifically, the belief of the actual configuration of the self (in orange or cyan), the belief of the other in relation to the self (in dark red or dark blue), and the belief of the other (in green or purple). Note the belief succession during goal-directed behavior: the potential configuration pointing toward one’s intentional state (either the elbow or the hand), followed by the current belief, and then by the real arm. Also note a slight delay about the inference of the configuration of the other agent.

limiting the agent to realize, e.g., multi-step reaching movements through fixed positions. In [65], a hybrid model in which the agent’s hypotheses were generated at each time step from the system dynamics allowed to relate continuous trajectories with discrete plans. Besides these more conventional solutions, many other hybrid methods have been proposed recently. One study addressed the problem of realistic robot navigation in active inference, making use of bio-inspired SLAM methods [106]. Other studies proposed how to integrate active inference with imitation learning for autonomous vehicles, using a Dynamic Bayesian Network (DBN) to explain the interactions between an expert agent and dynamic objects [107, 108]. An enhanced DBN consisting of two continuous levels and two discrete levels was used to model the behavior of a UAV at different timescales for assisting wireless communications [109, 110, 111, 112]. A study in robotics combined active inference with behavior trees for reactive action planning in dynamic environments [113]. Finally, a hybrid model based on recurrent switching linear dynamical systems allowed to discover non-grid discretizations of a continuous Mountain Car task. [114].

4.1 Dynamic inference by model reduction

Recall that in the unit of Figure 5a, some sort of multi-step behavior was achieved, based on higher-level priors over tactile sensations. In most cases, we need to switch intentions based on lower-level information, affording a more dynamic behavior. Considering a pick-and-place operation, an IE module would be more confident about the success of the first reaching movement if it could rely not just on a tactile belief but also on its kinematic configuration. In other words, hidden causes v should manage to effectively use both its prior η_v and the dynamics prediction error ε_x . The latter assumes two different roles depending on which pathway it flows into: the gradient with respect to the hidden states infers the position that is most likely to have generated the current trajectory, needed for movement; conversely, the gradient with respect to the hidden causes infers the most likely combination of gains v_m , signaling the current status of the trajectory and resulting in a dynamic modulation of intentions. However, since the hidden causes are generated by Gaussian distributions, they do not encode proper probabilities, and the gradient $\partial_v f_x$ infers just one over many possible combinations of gains. Instead, what we need is to infer *the most likely intention* to have generated the

current trajectory. Thus, to implement a correct intention selection, we assume that the hidden causes are generated from a categorical distribution:

$$p(\mathbf{v}) = \text{Cat}(\mathbf{H}_v) \quad (31)$$

where \mathbf{H}_v is a prior preference like $\boldsymbol{\eta}_v$. In this way, each discrete element of \mathbf{v} represents the probability that a specific continuous trajectory will be realized.

Conversion between discrete hidden causes to continuous hidden states (and vice versa) is done via Bayesian model reduction, a technique used to constrain the complexity of full posterior models into simpler and more restrictive (formally called *reduced*) distributions [100, 101]. Reduced means that the likelihood of some data is equal to that of the full model and the only difference rests upon the specification of the priors; hence, the posterior of a reduced model m can be expressed in terms of the posterior of the full model:

$$p(\tilde{\mathbf{x}}|\mathbf{o}, m) = p(\tilde{\mathbf{x}}|\mathbf{o}) \frac{p(\tilde{\mathbf{x}}|m)p(\mathbf{o})}{p(\tilde{\mathbf{x}})p(\mathbf{o}|m)} \quad (32)$$

In our case, model reduction means to explain the infinite values of a continuous signal with a discrete set of hypotheses. A simplified version of a hybrid active inference model is shown in Figure 13a. We can cast this procedure into the usual message passing, where top-down and bottom-up messages between the two domains perform a Bayesian Model Average (BMA) of reduced priors and a Bayesian Model Comparison (BMC) of reduced sensory evidence, respectively. In conventional hybrid models, discrete hidden states generate priors for the continuous hidden causes by weighting a specific reduced prior with the probability of each discrete state; hence, the reduced priors represent alternative hypotheses over the true causes of the sensorium [49]. The posterior over the hidden causes is then compared with such reduced priors to find which one among them could be the best explanation of the environment, taking into account their discrete probabilities before observing sensory evidence.

Because the agent compares continuous alternatives that are fixed and determined a-priori, it cannot correctly operate in a changing environment. For instance, if the agent thinks to find an object in one of two locations, it will always reach either one or the other initial guesses, even if the object has been moved to a third location. How then to use the newly available evidence to update our reduced assumptions? By considering the hidden causes as generated from a categorical distribution – as in Equation 31 – we can compare the posterior over the hidden states with the output of the dynamics functions \mathbf{f}_m , which thus act as the agent’s reduced priors [65]. More formally, we define M reduced prior probability distributions and a full prior model:

$$\begin{aligned} p(\mathbf{x}'|\mathbf{x}, m) &= \mathcal{N}(\mathbf{f}_m(\mathbf{x}), \mathbf{\Pi}_{x,m}) \\ p(\mathbf{x}'|\mathbf{x}, \mathbf{v}) &= \mathcal{N}(\boldsymbol{\eta}'_x, \mathbf{\Pi}_x) \end{aligned} \quad (33)$$

where $\boldsymbol{\eta}'_x$ is the full prior. Note here that the reduced priors have the same form of Equation 15 but are not directly conditioned on the hidden causes:

$$\mathbf{f}_m(\mathbf{x}) = \mathbf{e}_{i,m} = \mathbf{i}_m(\mathbf{x}) - \mathbf{x} \quad (34)$$

Next, we define the corresponding posterior models:

$$\begin{aligned} q(\mathbf{x}'|m) &= \mathcal{N}(\boldsymbol{\mu}'_{x,m}, \mathbf{P}_{x,m}) \\ q(\mathbf{x}') &= \mathcal{N}(\boldsymbol{\mu}'_x, \mathbf{P}_x) \end{aligned} \quad (35)$$

Now, we can find the full prior and its prediction error by averaging the continuous trajectories with their respective discrete probabilities:

$$\begin{aligned} \boldsymbol{\eta}'_x &= \sum_m v_m \mathbf{f}_m(\mathbf{x}) = \sum_m v_m \mathbf{e}_{i,m} \\ \boldsymbol{\varepsilon}_x &= \boldsymbol{\mu}'_x - \boldsymbol{\eta}'_x \end{aligned} \quad (36)$$

which have the same form of Equations 15 and 17. In fact, the hidden states still perceive a single dynamics prediction error containing the total contribution of every intention. Concerning the bottom-up messages \mathbf{l} , we first write the free energy of each reduced model in terms of the full model. As before, maximizing each reduced free energy makes it

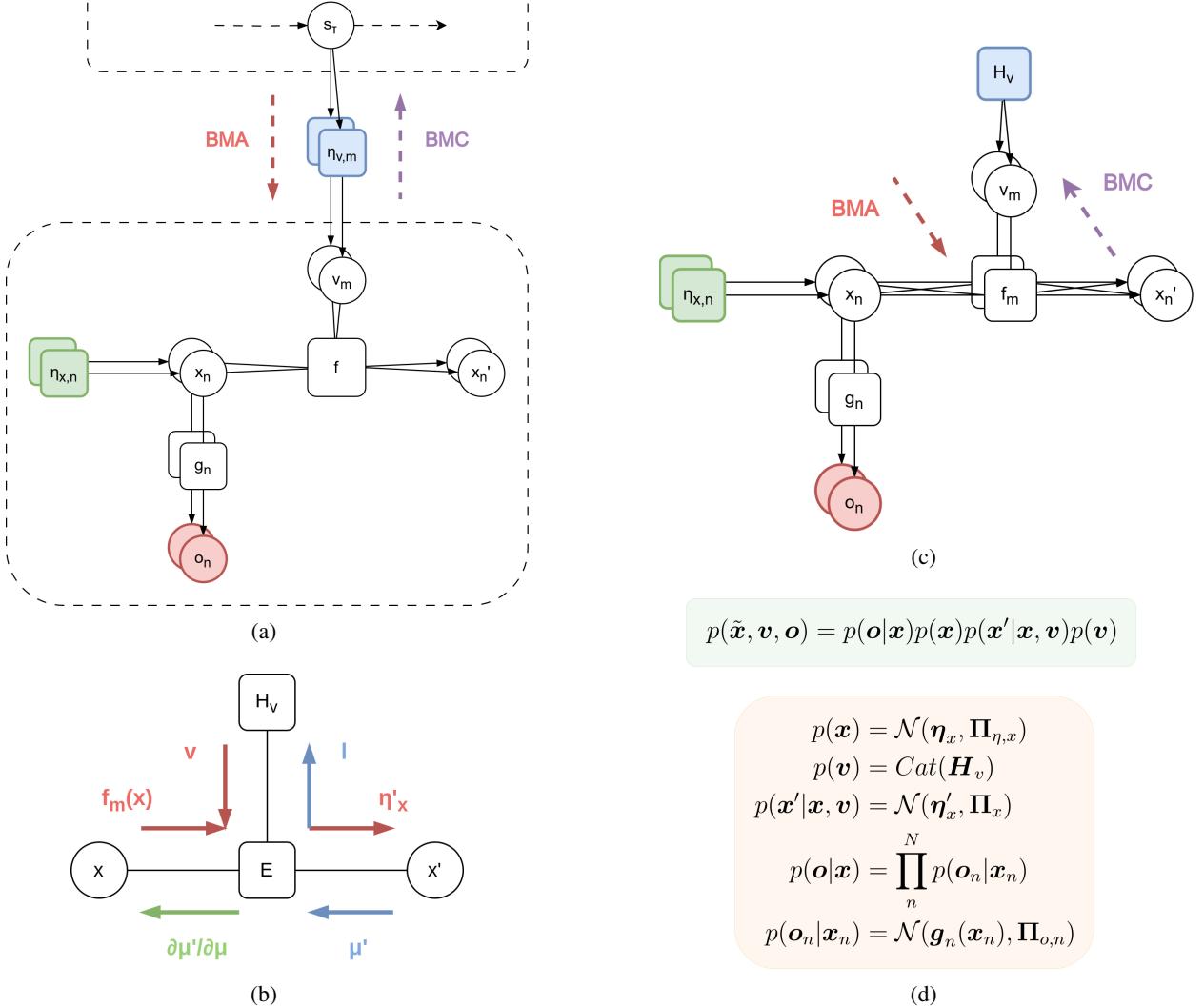


Figure 13: (a) A conventional hybrid model, composed of a discrete model at the top and a continuous model at the bottom. For simplicity, we assume that the continuous prior η_v is directly conditioned on discrete hidden states s . We will cover discrete models in the following section. Here, top-down and bottom-up messages are computed by Bayesian model reduction of some static priors, without the possibility for dynamic planning. (b) A simplified graph displaying the exchange of top-down (red) and bottom-up (blue and green) messages of the hybrid control in (c). (c) Factor graph of the unit with hybrid control. The hidden causes are now generated from a categorical distribution, such that instead of inferring a combination of continuous attractor gains, the model correctly infers the most likely intention associated with the current trajectory. This is done by computing the free energy E_m corresponding to each intention. Through Bayesian model reduction between discrete hidden causes and continuous hidden states, the agent updates its reduced priors at each time step. Generative model of (c).

approximate the log evidence:

$$\mathcal{F}(m) = \mathcal{F} - \ln \int \frac{p(\tilde{\mathbf{x}}|m)}{p(\tilde{\mathbf{x}})} q(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} \approx \ln p(\mathbf{o}|m) \quad (37)$$

As a result, the free energy related to each dynamics function f_m depends on the approximate posterior $q(\tilde{\mathbf{x}})$ of the full model, avoiding the computation of the reduced posteriors. Under a Gaussian approximation, the m th reduced free energy breaks down to a simple formula and the bottom-up messages l are found by accumulating the log evidence

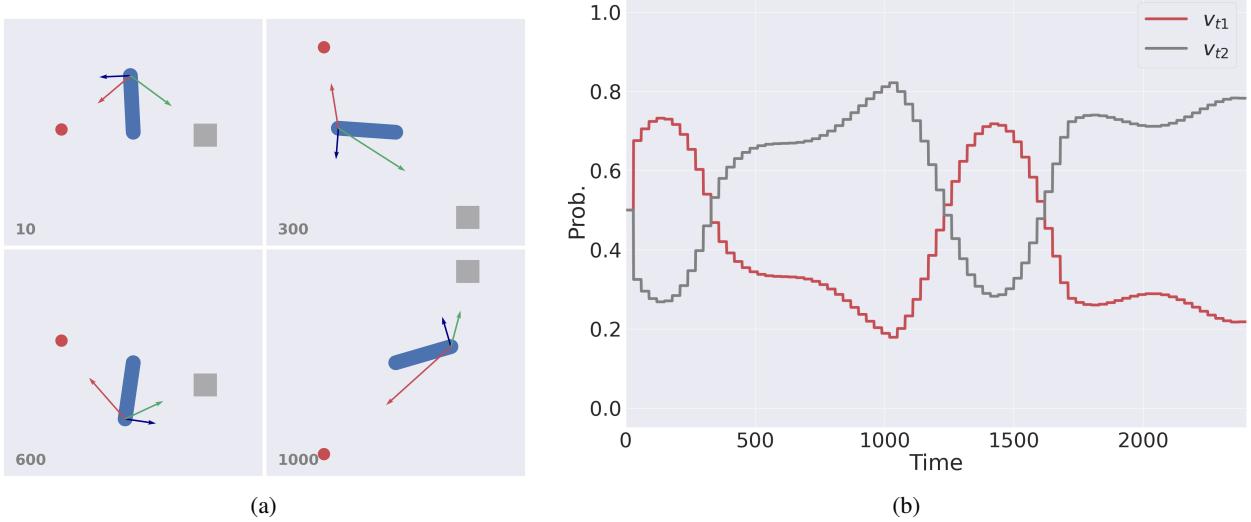


Figure 14: (a) In this perception task, the agent has to infer which one among two objects (a red ball or a grey square) is following. The arm moves along a circular trajectory, while the two objects move linearly in opposite directions. (a) Sequence of time frames (left). The dark blue arrow represents the real agent’s trajectory, while the red and green arrows represent potential trajectories associated with different reaching movements. (b) Dynamics of hidden causes. The two hidden causes v_{t1} and v_{t2} are associated with the two potential trajectories. As the hand moves away from the first target and approaches the second target, the dynamic accumulation of evidence leads to an increase in the second hidden cause. Here, we used a uniform prior \mathbf{H}_v , so the hidden causes only express the accumulated log evidence \mathbf{l} . Also, the continuous time T for evidence accumulation was set to 30 time steps. See [65] for more details.

associated with every intention for a certain amount of continuous time T :

$$l_m = \int_0^T \mathcal{L}_m dt \quad (38)$$

$$\mathcal{L}_m = \frac{1}{2} (\boldsymbol{\mu}_{x,m}'^T \mathbf{P}_{x,m} \boldsymbol{\mu}_{x,m}' - \mathbf{f}_m(\mathbf{x})^T \boldsymbol{\Pi}_{x,m} \mathbf{f}_m(\mathbf{x}) - \boldsymbol{\mu}_x'^T \mathbf{P}_x \boldsymbol{\mu}_x' + \boldsymbol{\eta}_x'^T \boldsymbol{\Pi}_x \boldsymbol{\eta}_x')$$

Then, a BMC turns into computing the softmax of a vector comprising the free energy E_m of every reduced model. This quantity compares the prior surprise $-\ln \mathbf{H}_v$ with the accumulated log evidence:

$$v = \sigma(-E) = \sigma(\ln \mathbf{H}_v + \mathbf{l}) \quad (39)$$

where $\mathbf{l} = [l_1, \dots, l_M]$. See [100, 101] for a full derivation of BMC under the Laplace assumption, and [65] for more details about the presented approach. Equation 39 is the discrete analogous of Equation 9, but now the bottom-up message encodes a proper discrete distribution and can be used to infer the most likely intention associated with the current trajectory.

The factor graph of this model, which we call a *hybrid unit*, is displayed in Figure 13c. Its inference process at each continuous step is better understood if we analyze separately the three different pathways shown in Figure 13b: (i) during the forward pass, the unit receives a discrete intention prior \mathbf{H}_v , performs a BMA with potential trajectories $\mathbf{f}_m(\mathbf{x})$, and imposes a prior $\boldsymbol{\eta}_x'$ over the 1st order; (ii) through the first backward pass, the unit accumulates the most likely intention related to the current trajectory by comparing it to the ones generated by the dynamics functions; (iii) in the second backward pass, the unit propagates the dynamics prediction error back to the 0th order to infer the most likely continuous state associated with the trajectory, eventually generating biased observations. After a period T , the unit finally computes the difference between the discrete prior and the accumulated evidence, generates a new combination of intentions, and the process starts over.

Now how this mechanism is self-reinforcing: a decision generates some movement, which in turn infers the decision itself to compute the next movement. Inferring a target from body movements is a fundamental aspect often neglected in decision-making studies, which can result in completely different behaviors. In particular, this component produces a

commitment effect to the decision taken and stabilizes the actions, avoiding changes of mind that may lead to the loss of valid opportunities in dynamic environments [115, 116]. Further, this kind of dynamic inference has several utilities, e.g., it can be used to infer which one among multiple objects an agent is following – as exemplified in Figure 14 – by generating trajectories for different objects and comparing them with the one perceived [65].

An alternative that produces similar results is the model used in [117] in the context of social exchange. In this hybrid solution, a "student" bird maintains several continuous generative models for every teacher conspecifics. Discrete switching variables placed at the center of these hypotheses infer which teacher bird has generated the perceived birdsong. Learning of the generative models relies on two complementary methods, i.e., Bayesian model average of all possible teacher birds, or Bayesian model selection of a specific bird generating a song.

As a last note, the dynamics precisions of Equations 33 and 38 have here an interesting interpretation equivalent to the observation precisions Π_o . Predictive coding assumes that whenever an agent perceives high noise about a sensory modality, the precision of that generative model will decrease because it cannot be trusted for understanding the state of affairs of the world [19, 20]. In addition, the dualism between action and perception inherent to the free energy principle tells us that the optimization of precisions – which are thought to be encoded as synaptic gains – could play a crucial role in attention mechanisms that selectively sample sensory data [118, 119]. Based on these assumptions, we note a dual interpretation of (reduced) dynamics precisions. A low precision $\Pi_{x,m}$ related to a grasping action could mean that it is unreliable to explain the current context (e.g., an object far away from the hand). In addition, it could mean that the agent does not intend to rely on it to achieve a goal (e.g., grasping an object when it is out of reach). This perspective unveils an additional mechanism besides the fast inference of hidden causes that we mentioned before: a slow learning of reduced precisions that lets the agent score – and, crucially, focus on – those intentions that would be appropriate for a specific scenario [65].

4.2 A discrete interface for dynamic planning

Numerous studies have demonstrated that the brains of athletes are marked by a higher activation of posterior and subcortical regions that involves little or no conscious thinking, producing fluid transitions between different motions; in contrast, the brain of a novice requires a higher demand of prefrontal computations that results in lower performances [120, 121, 122]. From an active inference perspective, we can compare the proficiency of athletes with the continuous model of Figure 5a, corresponding to the subcortical sensorimotor loops. This model encodes a transition mechanism that is not very flexible, but that precisely for this can react much more rapidly to environmental stimuli, e.g., when grasping objects moving at high speed [66]. In general, this strategy can be very effective when the environment has limited uncertainty and the task to be solved comprises a rigid sequence of actions, which the agent has already correctly learned. However, suppose that the agent is introduced to a novel or complex task that requires careful thinking about the imminent future. In this case, it should be capable of replanning the correct sequence of actions if something goes unexpectedly, and a high-level belief always producing an a-priori-determined behavior for the hidden causes would fail to complete the task.

Having replaced the continuous hidden causes of Figure 5a with discrete hidden causes in Figure 13c, we can now endow the agent with planning capabilities through a discrete model composed of the following distributions – as shown in Figure 15:

$$p(\mathbf{s}_{1:\tau}, \mathbf{v}_{1:\tau}, \boldsymbol{\pi}) = p(\mathbf{s}_1)p(\boldsymbol{\pi}) \prod_{\tau}^{\tau} p(\mathbf{v}_{\tau}|\mathbf{s}_{\tau})p(\mathbf{s}_{\tau}|\mathbf{s}_{\tau-1}, \boldsymbol{\pi}) \quad (40)$$

where:

$$\begin{aligned} p(\mathbf{s}_1) &= \text{Cat}(\mathbf{D}) & p(\mathbf{v}_{\tau}|\mathbf{s}_{\tau}) &= \text{Cat}(\mathbf{v}_{\tau}|\mathbf{A}\mathbf{s}_{\tau}) \\ p(\boldsymbol{\pi}) &= \sigma(-\mathcal{G}) & p(\mathbf{s}_{\tau}|\mathbf{s}_{\tau-1}, \boldsymbol{\pi}) &= \text{Cat}(\mathbf{s}_{\tau}|\mathbf{B}_{\boldsymbol{\pi}, \tau}\mathbf{s}_{\tau-1}) \end{aligned} \quad (41)$$

Here, \mathbf{A} , \mathbf{B} , \mathbf{D} are the likelihood matrix, transition matrix, and prior of the hidden states, $\boldsymbol{\pi}$ are the policies (which are not state-action mappings as in RL but sequences of actions), \mathbf{s}_{τ} are the discrete hidden states at time τ . These quantities have strict analogies with their continuous counterparts of Figure 1b, i.e., the likelihood function \mathbf{g} , the dynamics function \mathbf{f} , the prior η_x , hidden causes \mathbf{v} , and hidden states \tilde{x} , with the difference that the discrete hidden states do not encode instantaneous paths expressed in generalized coordinates, but sequences of future states. Crucially, \mathcal{G} is the *expected free energy*, defined as the free energy that the agent expects to perceive in the future.

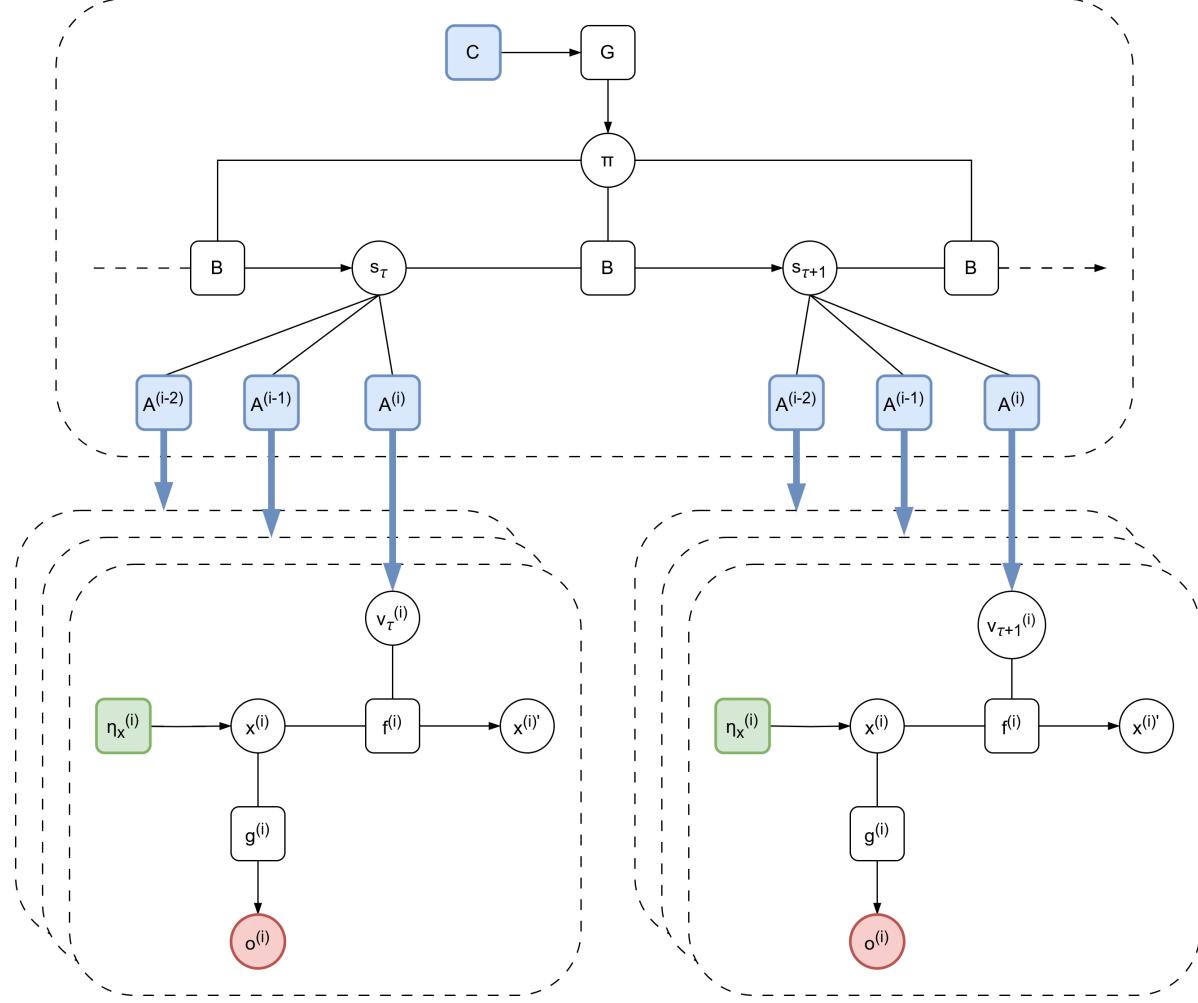


Figure 15: Interface between a discrete model (at the top) and several hybrid units (at the bottom). For clarity, the hidden states factorization of each unit is not displayed. The discrete hidden states s at time $\tau + 1$ are computed from the current hidden states s_τ , by choosing some policy π_π related to a specific transition distribution encoded in B . The best policy at any moment is the sequence of actions that most minimizes the expected free energy \mathcal{G} . As a result, the agent will try to sample those observations that conform to its preference C . The discrete hidden causes $v^{(i)}$ at time τ are directly generated from discrete hidden states s_τ through likelihood matrices $A^{(i)}$, thus affording dynamic planning, synchronized behavior, and inference with multiple evidences.

Note here that the likelihood matrix A expresses a conditional probability over the discrete hidden causes v_τ . As in conventional hybrid models, the discrete hidden states are linked to the hidden causes, but now the latter directly act as discrete outcomes generated by the likelihood matrix, which thus replaces the prior H_v in Equation 31. Here, we wish to infer the posterior distribution:

$$p(s_{1:\tau}, \pi | v_{1:\tau}) = \frac{p(v_{1:\tau} | s_{1:\tau}, \pi)p(s_{1:\tau}, \pi)}{p(v_{1:\tau})} \quad (42)$$

As before, this requires computing the intractable model evidence $p(v_{1:\tau})$, so we resort to a variational approach: we first express the approximate posterior by its sufficient statistics $s_{\pi,\tau}$ and conditioning upon a specific policy:

$$p(s_{1:\tau} | v_{1:\tau}, \pi) \approx q(s_{1:\tau}, \pi) = q(\pi) \prod_{\tau}^{\tau} q(s_\tau | \pi) \quad (43)$$

where:

$$\begin{aligned} q(\boldsymbol{\pi}) &= \text{Cat}(\boldsymbol{\pi}) \\ q(\mathbf{s}_\tau | \boldsymbol{\pi}) &= \text{Cat}(\mathbf{s}_{\pi,\tau}) \end{aligned} \quad (44)$$

We then compute the variational free energy \mathcal{F}_π of that policy:

$$\mathcal{F}_\pi = \sum_{\tau=1}^T \mathbf{s}_{\pi,\tau} \ln \mathbf{s}_{\pi,\tau} - \sum_{\tau=1}^T \mathbf{v}_\tau \ln \mathbf{A} \mathbf{s}_{\pi,\tau} - \mathbf{s}_{\pi,1} \ln \mathbf{D} - \sum_{\tau=2}^T \mathbf{s}_{\pi,\tau} \ln \mathbf{B}_{\pi,\tau} \mathbf{s}_{\pi,\tau-1} \quad (45)$$

Differentiating Equation 45 allows us to infer the most likely discrete hidden states at time τ for policy π :

$$\mathbf{s}_{\pi,\tau} = \sigma(\ln \mathbf{B}_{\pi,\tau-1} \mathbf{s}_{\pi,\tau-1} + \ln \mathbf{B}_{\pi,\tau}^T \mathbf{s}_{\pi,\tau+1} + \ln \mathbf{A}^T \mathbf{v}_\tau) \quad (46)$$

where we applied a softmax function to ensure that it is a proper probability distribution. In order to infer the policies, we additionally consider unobserved outcomes as random variables, and then condition the joint probability of the generative model upon some preferences \mathbf{C} . In this way, the agent can predict future outcomes and select the most likely sequence of actions that will lead it to its desired state. The expected free energy \mathcal{G}_π under policy π consists of a *pragmatic* or goal-seeking term toward the agent's preferences, and an *epistemic* or uncertainty-reducing term (see [96] for more details):

$$\begin{aligned} \mathcal{G}_\pi &\approx \sum_{\tau} D_{KL}[q(\mathbf{v}_\tau | \boldsymbol{\pi}) || p(\mathbf{v}_\tau | \mathbf{C})] + \underset{q(\mathbf{s}_\tau | \mathbf{s}_{\tau-1}, \boldsymbol{\pi})}{\mathbb{E}} [H[p(\mathbf{v}_\tau | \mathbf{s}_\tau)]] \\ &= \sum_{\tau} \mathbf{v}_{\pi,\tau} (\ln \mathbf{v}_{\pi,\tau} - \mathbf{C}_\tau) + \mathbf{s}_{\pi,\tau} \mathbf{H}_A \end{aligned} \quad (47)$$

where:

$$\mathbf{v}_{\pi,\tau} = \mathbf{A} \mathbf{s}_{\pi,\tau} \quad \mathbf{C}_\tau = \ln p(\mathbf{v}_\tau | \mathbf{C}) \quad \mathbf{H}_A = -\text{diag}(\mathbf{A}^T \ln \mathbf{A}) \quad (48)$$

Here, $\mathbf{v}_{\pi,\tau}$ is the prediction of the hidden causes at time τ and under policy π , \mathbf{C}_τ is the logarithm of the prior preference over the hidden causes, and \mathbf{H}_A is the ambiguity of the likelihood matrix \mathbf{A} . The policy π is then found by applying a softmax function to the expected free energy, as defined in Equation 41.

Using the (deep) hierarchical model outlined in the previous section, we may connect several hybrid units to the discrete model by appropriate likelihood matrices $\mathbf{A}^{(i)}$. Each of them has an independent interface whereby the discrete model computes different signals and waits for the next step $\tau + 1$, when it can infer its hidden states based on multiple accumulated evidences. Recall that in the combined structure of Figure 15, the role of the hybrid unit was to predict a trajectory from a discrete intention prior, and to infer the most likely intention in a continuous period T . But now the intention prior is generated from the discrete model:

$$\mathbf{v}_\tau = \sigma(\ln \mathbf{A} \mathbf{s}_\tau + \mathbf{l}_\tau) \quad (49)$$

where \mathbf{l}_τ is the bottom-up message at time τ found by Equation 38, and we compute \mathbf{s}_τ by marginalizing over all policies, i.e. $\mathbf{s}_\tau = \sum_\pi \pi_\pi \mathbf{s}_{\pi,\tau}$.

In sum, computing the posterior probability over policies π turns into finding the best action that conforms to the dual objective defined by \mathcal{G} . Here, the discrete actions are not intended as actual motor commands similar to Equation 7, but as abstract actions over high-level representations. In fact, the hierarchical nature of discrete models in active inference permits performing decision-making with a separation of temporal scales, wherein a specific level can generate and infer the states and the paths of the level below [123, 124, 125]. Further evaluating the consequences of an action for a longer time horizon affords more advanced planning called *sophisticated inference* [126]. Computing actions with the expected free energy is different from the motor control of continuous models, which only minimize the free energy of present states.

In addition to the previous agents, it is now possible to synchronize the behavior of different continuous signals based on the same high-level policy. For instance, one can realize a pick-and-place operation with a moving object – as represented in Figure 16 – producing smooth transitions between reaching and grasping actions, respectively performed

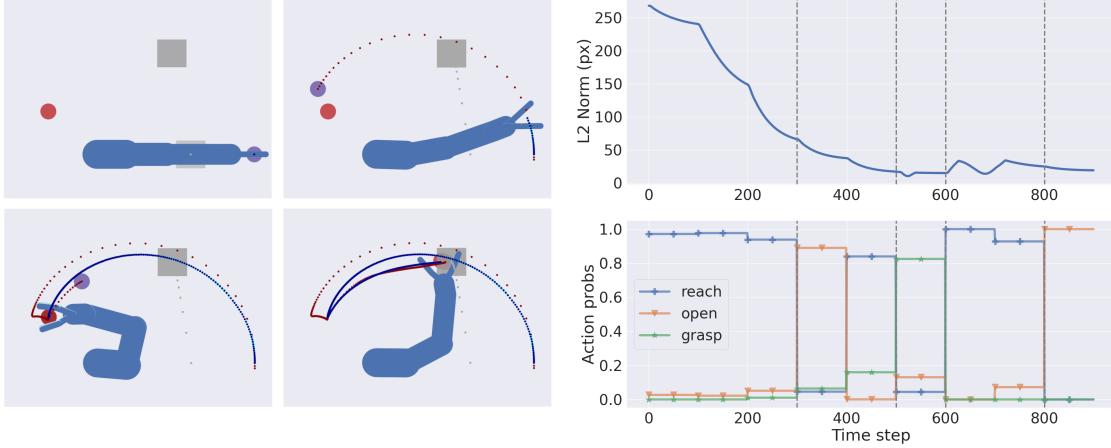


Figure 16: In this task, the agent (a 4-DoF arm with an additional 4-DoF hand composed of two fingers) has to pick a moving object, and place it at a goal position. The agent has a shallow structure with a single IE module computing the hand position from the arm joint angles. Besides the spatial dimension, the agent can easily integrate multisensory information such as touch, dimensions of objects, and other more abstract properties. Note that the object belief is rapidly inferred, and as soon as the picking action is complete, the belief is gradually pulled toward the goal position, resulting in a second reaching movement. The top right panel shows the hand-object distance over time, while the bottom right panel displays the dynamics of the discrete action probabilities used to infer the next discrete state. See [66] for more details.

in extrinsic and intrinsic domains [66]. Note how an intermediate phase between the two actions naturally arises, corresponding to a composite approaching movement. In principle, the learning of dynamics precisions $\Pi_{x,m}$ might shed light on how motor skill learning occurs, via message passing between continuous trajectories and discrete policies. Moreover, through this kind of dynamic planning the agent can infer and realize instantaneous trajectories even for the same continuous period T within a discrete step τ , useful, e.g., for grasping the moving object without waiting for the successive replanning step. Third, this configuration permits learning a discrete representation of the environment based on continuous evidence: as shown in [127], learning the likelihood matrix A involves counting the coincidences between targets and body movements, thus adapting a response strategy (risky versus conservative) depending on the difficulty of the context. Additionally, through learning of the prior D , a habitual behavior toward the chosen decision emerges. Similarly, one could update preferred states encoded in the matrix C based on current observations and actions.

Since the agent plans with trajectories and not positions, to correctly maintain a goal state we need to introduce a hidden cause loosely corresponding to the *stay* action commonly used in discrete tasks [97]. This hidden cause can be linked to an identity map, i.e., $i_{\text{stay}}(\mathbf{x}) = \mathbf{x}$, which can be interpreted as the agent’s desire to maintain the current state of the world [65]. This hidden cause also relates to the initial state of the task, and translates into a phase of pure perceptual inference.

4.3 Flexible hierarchies

Figure 17 portrays a *deep hybrid model* designed for solving a tool-use task [90]. It combines the expressivity of a (deep) hierarchical formulation, the advantages of planning trajectories inherent to a hybrid unit, and the possibility of encoding object affordances and other agents. As in Figure 15, the IE modules communicate with a discrete model at the top, but now they are combined in a hierarchical fashion recapitulating the agent’s kinematic chain. As a consequence, two different goal-directed strategies arise. Considering a simple reaching movement, an attractor imposed at the hand level would generate a cascade of extrinsic prediction errors flowing back to the previous levels and finding a suitable kinematic configuration with the hand over the target. This corresponds to a *horizontal* hierarchical depth occurring along the hybrid units, and can be compared to the process of motor babbling typical of infants [128], whereby random attractors are generated at different hierarchical levels to identify the correct body structure. In addition to this naive strategy, since a discrete model can now generate trajectories for every IE module (in both intrinsic and extrinsic domains), a more advanced behavior can be achieved once inverse kinematics is correctly performed, which imposes a

specific path to the whole kinematic chain. This corresponds to a *vertical* hierarchical depth with two (discrete and continuous) temporal scales, steering the lower-level inference in a direction that, e.g., avoids singularities or gets out from local minima generated by repulsive attractors.

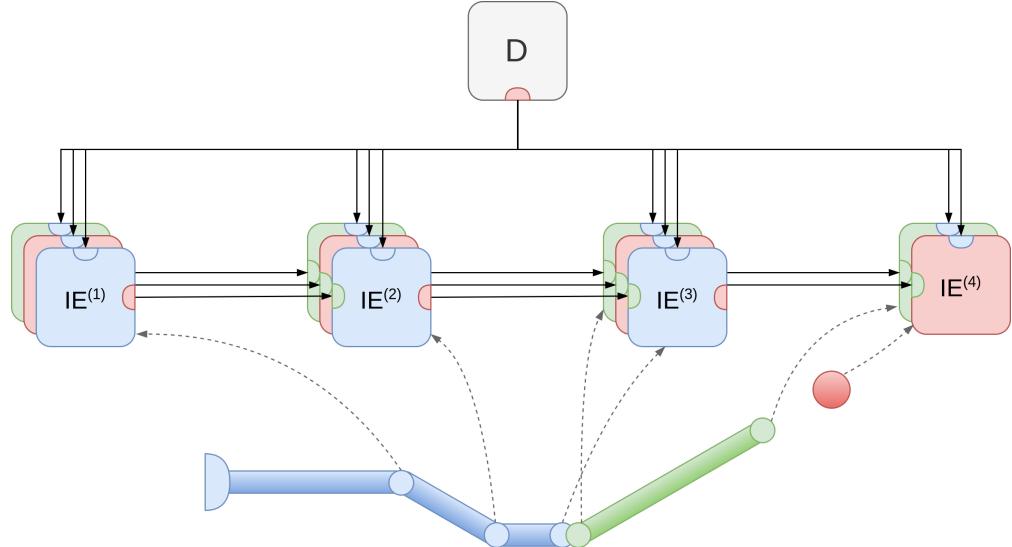


Figure 17: Graphical representation of a deep hybrid model for tool use, composed of a discrete model at the top and several IE modules. Every module is factorized into three elements, which are respectively linked to the observations of the agent’s arm (in blue), a tool (in green), and a ball (in red). Note that the last level only encodes the tool’s extremity and the ball.

Considering the update rule for the (extrinsic) hybrid unit at level i :

$$\dot{\mu}_{x,e}^{(i)} \propto -\Pi_{o,e}^{(i-1)} \varepsilon_{o,e}^{(i-1)} + \partial g_e^T \Pi_{o,e}^{(i)} \varepsilon_{o,e}^{(i)} + \partial \eta_{x,e}^{(i)T} \Pi_{x,e}^{(i)} \varepsilon_{x,e}^{(i)} \quad (50)$$

where $\partial \eta_{x,e}^{(i)T}$ is the gradient of the trajectory prior of Equation 36, we note a delicate balance between forward and backward extrinsic likelihood, and the top-down modulation of the discrete model. From the discrete model’s perspective, the discrete hidden states produce a specific combination of hidden causes for each hybrid unit; this combination generates a composite trajectory in the continuous domain weighting distinct potential trajectories, taking into account dynamic elements for the whole discrete step τ . After this period, evidence is accumulated for every hybrid unit, eventually inferring the most likely discrete state that may have generated the actual trajectory, related to the self and the environment.

A non-trivial issue exists in tasks requiring tool use, e.g., reaching a ball with the extremity of a stick. Much as other agents may have different kinematic structures than the self, a tool may have its own hierarchy (e.g., even a simple stick is represented by two Cartesian positions and an angle) that must somehow be integrated into the agent’s generative model. Specifically, reaching an object with the extremity of a tool means defining a potential body configuration augmented by a new *virtual* level. This new level does not exist in the generative process and the belief of the actual body configuration x_0 . Nonetheless, if we consider the two potential configurations, the agent thinks of the tool as an extension of its arm, hence flexibly modifying its body schema as discussed in the Introduction. This is possible by linking the two visual observations of the tool to the hand and virtual levels in a second pathway of the hidden states, as shown in Figure 17. Since the intrinsic units of the IE modules also encode information about limb lengths, the agent can not only infer its kinematic structure through visual observations, but also the actual length of the tool [129]. While this second pathway is still marked by a clear distinction between the tool and the arm since the hand level receives observations from both elements, a third pathway is constructed such that the observation of the ball is only linked to the virtual level. As a result, this new potential configuration views the arm and the tool as part of the same kinematic chain. The interactions between these three pathways (shown in Figure 18) may shed light on how the remapping of the motor cortex gradually occurs with extensive tool use [10, 11], modifying the boundaries between the self and the environment.

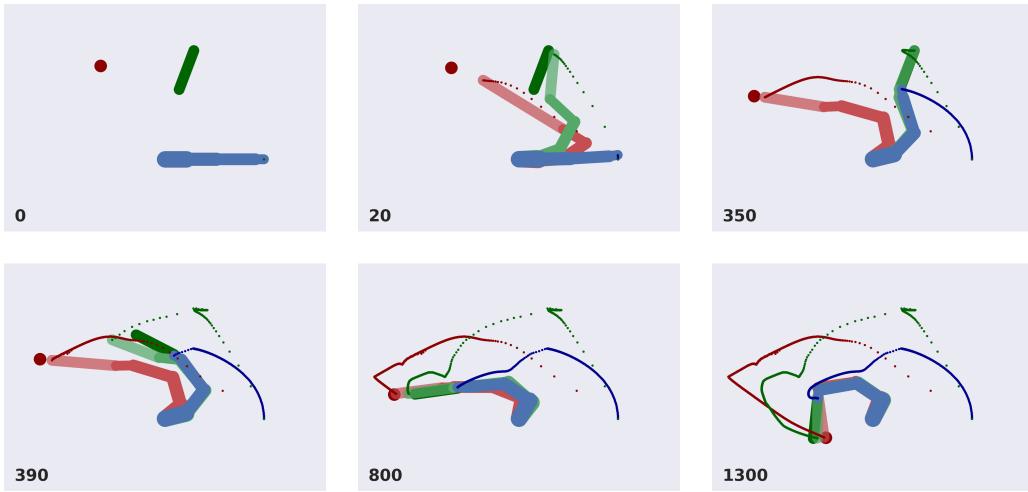


Figure 18: In this task, the agent (a 4-DoF arm) has to grasp a green tool and track a moving red ball with the tool’s extremity. The real arm configuration is represented in blue, while the light green and light red arms correspond to the potential agent’s configurations in relation to the tool and the ball, respectively. Note that the last two levels of the tool’s belief gradually match the real tool, while the ball’s belief makes no distinction between the arm and the tool, and is only defined by the tool’s extremity. The (deep) hierarchical factorization allows the agent to infer a potential configuration for the ball even during the first reaching movement. See [90] for more details.

5 Discussion

Despite the many advances that have been made in active inference, with increasing popularity among different scientific domains, a current drawback is that studies about motor control and decision-making have been somewhat separated so far, making use of two highly similar but distinct frameworks. As a result, there is no consensus on how to achieve dynamic planning (i.e., how to perform decision-making in constantly changing environments), and state-of-the-art solutions to tackle complex tasks generally couple active inference with traditional methods in machine learning or optimal control. From a theoretical perspective, a few works prescribed an efficient and elegant way for combining the capabilities of discrete and continuous representations into a single generative model [49, 67]; however, this hybrid approach has not reached as much maturity, with the consequence that there are far fewer studies on the subject in the literature.

For this reason, we tried here to give a comprehensive view of this yet unexplored direction, comparing several design choices regarding tasks of increasing complexity, with the intent of bringing motor control and behavioral studies closer. As a practical example, we described the modeling of tool use [90], a task that inevitably calls for both discrete and continuous frameworks, and that requires taking two additional aspects into account, i.e., object affordances and hierarchical relationships. In a simple scenario, considering a target to reach as the cause of some hidden states is a reasonable assumption and permits operating in dynamic contexts. But when there are multiple objects, how does the agent decide what the cause of a particular action will be? How can it account for different object affordances? And what if the target moves along a non-trivial path? The hidden states may be factorized into independent distributions encoding multiple entities in intrinsic coordinates, hence expressing *potential body configurations*. Further, the hidden causes may be linked to *potential trajectories* related to the agent’s intentions [64, 80, 66]. Beliefs over each entity would have their own dynamics, allowing the agent to predict, e.g., the trajectory of a moving ball. Then, we described how this unit could be scaled up to construct complex (deep) hierarchical structures, e.g., for simulating human body kinematics [80], and to perform more general transformations of reference frames, e.g., perspective projections [83]. A hierarchical factorization of the hidden states now assumes a broader perspective that can also account for multi-agent interaction – an aspect that has been analyzed in the discrete framework as well [130]. Finally, we described the design of a hybrid unit with discrete hidden causes and continuous hidden states, affording dynamic inference via Bayesian model reduction [65]; then, a higher-level discrete model made it possible to simulate dynamic tasks involving online

planning of actions. This showed further parallelisms between the inference of trajectories in continuous models and policies in discrete models.

One challenge in maintaining a deep representation of the kinematic chain is the associated computational complexity and the time required to infer a body posture from visual input, both increasing with the DoF. This is because an extrinsic prediction error generated at the distal (e.g., hand) levels would have to climb the hierarchy toward the (root) body-centered reference frame. This directly affects behavioral accuracy and movement time, which critically depend on the correct inference of the intrinsic state. A performance comparison during inference and action is shown in [80], for increasing DoF. In contrast, handling multiple objects does not lead to a significant increase in inference time, since the (deep) representations are computed in parallel, and are only limited by the complexity of the object dynamics. Communication with the discrete model is equally efficient, as a single discrete state can infer the body trajectories based on multiple units concurrently [65]. Still, the behavior of other objects in the environment requires predicting not only how the agent will behave after a specific action but also how the other objects will behave and their influence on the agent's behavior. This requires a rich discrete representation that can model the interactions of every object, and evaluating complex policies may not be as effective for high-dimensional settings [131]. A last issue regards the time T needed for accumulating continuous evidence for a single discrete step τ . As shown in [66], Figure 5b, a narrow sampling time interval permits efficient control in highly dynamic environments but at the cost of increased computational time.

A limitation of the models reviewed here is their fixed structure. The critical question, then, is how tool use can be realized without embedding prior knowledge into the agent's generative model. In other words, how an agent can perform using active inference when starting from a blank memory or under the assumption that the surrounding environment will remain unchanged. A common criticism of continuous-time active inference models is that their generative models are a-priori defined and fixed, with intricate and hardcoded dynamics functions that raise some concerns about biological plausibility. In contrast, one appealing characteristic of PCNs is that they simulate brain processing with extremely simple functions typical of the connectivity of neural networks (e.g., linear combinations of weights and biases passed to a non-linear activation function). This allows PCNs to easily adapt to high-dimensional data, with a few critical advantages compared to deep learning (e.g., about top-down modulation) [25]. While much of the research with PCNs involves static representations, some studies addressed how predictive coding could be used to learn temporal sequences [132, 89], or to solve RL tasks [133, 91, 134]. Here, we demonstrated how generative models in active inference could be realized by simple likelihood and dynamics functions, showing some analogies with the inference of PCNs. Based on these findings, a promising research direction would be to imitate their (deep) hierarchical architectures (as in Figure 9d), so that an agent could not only flexibly adapt its body schema to interact with objects with different hierarchical structures, but also learn the system dynamics and act over them to conform to prior beliefs.

Learning policies in continuous environments is not an easy challenge, but addressing it with strategies different from traditional methods might be key for advancing with current intelligent agents, realizing the full theoretical potential at the basis of active inference and the free energy principle. On this matter, the state-of-the-art is to approximate the likelihood and transition distributions by deep neural networks [135, 136, 137, 138, 139, 140, 141, 142]. While several benefits arise compared to deep RL, this still relegates the deep structure within the neural network, generally using a single-level active inference agent. One study used a more biologically plausible PCN as a generative model [134], but relied on a similar approach. As extensively analyzed in [57], neural networks can be seen as static generative models with infinitely precise priors at the last level and no hidden states. This architecture can be used to perform sparse coding or Principal Components Analysis (PCA); however, it fails to account for dynamic variables, as in deconvolution problems or filtering in state-space models. Temporal depth – either discrete or continuous – is thus key to inferring the most accurate representation of the environment, and indeed it seems that cortical columns are able to express model dynamics (e.g., the prefrontal cortex is constantly involved in predicting future states, and motion-sensitive neurons have been recorded in the early visual cortex as well [143]). While it is true that temporal sequences can be easily handled by deep architectures such as recurrent neural networks or transformers [144], their passive generative mechanism could still be reflected in the behavior of the active inference agent. In contrast to such a passive AI, being grounded on sensorimotor experiences and actively modifying the environment could be fundamental to the emergence of genuine understanding [145]. Taken together, these facts suggest that a hierarchy of actions upon generalized coordinates of motion or discrete future states could bring several advantages in solving RL tasks. For instance, representing an agent in a hierarchical fashion afforded highly advanced control over its whole body structure that would not have been possible by a single level generating only the hand position [80, 129].

How to learn dynamic planning in deep hierarchical models? The importance of being discrete when considering structure learning has been stressed in [146]. Indeed, hierarchical discrete models afford much more expressivity compared to their continuous counterparts, above all, deriving from the simplicity of computing the expected free energy. Nonetheless, as Friston and colleagues note, whether using continuous or discrete representations depends on the model evidence. Specifically, the former may have better performances when the evidence has contiguity properties, e.g., when dealing with time series or with Euclidean space. Indeed, the task exemplified in Figure 18 is effective because the Bayesian model reduction performs a dynamic evidence accumulation over the extrinsic space in which the agent operates. Hence, coupling the hierarchical depth of the hybrid units in Figure 17 with a hierarchical discrete architecture (and not just a single discrete level) could bring efficient structure learning also in constantly changing environments. A successful Bayesian approach to learning discrete structures involves growing discrete distributions using an infinite Dirichlet process and the Chinese restaurant prior [147]. This method assumes a potentially infinite mixture of basis distributions and builds a structure starting from an empty model wherein novel configurations are either assigned to popular existing states, or occasionally, used to create novel states. This approach has shown success in learning structures that support complex goal-directed behavior [148], hierarchical spatial organization [17], and spatial navigation [16] among other applications. An alternative to hierarchical discrete models would be to combine units composed of a joint discrete-continuous model – as in Figure 15 – which would allow to perform dynamic planning within each single unit. While this solution may not be supported by empirical evidence from biological agents, it could be an encouraging direction to explore from a machine learning perspective, contrasting the hypothesis of central discrete decision-making with a distributed network of local decisions.

A third interesting topic regards motor intentionality. Although multi-step tasks are typically tackled at the discrete level, we showed here that, under appropriate assumptions, a non-trivial behavior could be achieved and analyzed also at the continuous level. The flexible intentions we defined could be compared to an advanced stage of motor skill learning, consisting of autonomous and smooth movements that do not necessitate conscious decision-making [66]. Still, even in this case the model structure was predefined. How do such intentions emerge during repeated exposure to the same task? How does the agent score which intentions will be appropriate for a specific context? As mentioned in the last chapter, optimization of dynamics precisions is likely to involve the free energy of reduced models (see Equation 38). This process may shed light on how discrete actions arise from low-level continuous trajectories and, conversely, how the latter are generated from a composite discrete action. Last, a few studies proposed additional connections between policies unfolding at different timescales, either directly [124, 125] or through discrete hidden states [123]. Such approaches could be adopted in hybrid and continuous contexts as well, so that flexible intentions could be propagated via local message passing between hidden causes along the whole hierarchy.

Data availability

Code and data are deposited in GitHub (<https://github.com/priorelli/dynamic-planning>).

Acknowledgments

This research received funding from the European Union’s Horizon H2020-EIC-FETPROACT-2019 Programme for Research and Innovation under Grant Agreement 951910 to I.P.S. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- [1] Rossella Breveglieri, Claudio Galletti, Annalisa Bosco, Michela Gamberini, and Patrizia Fattori. Object affordance modulates visual responses in the macaque medial posterior parietal cortex. *Journal of Cognitive Neuroscience*, 27(7):1447–1455, July 2015.
- [2] Maria C. Romero, Pierpaolo Pani, and Peter Janssen. Coding of shape features in the macaque anterior intraparietal area. *The Journal of Neuroscience*, 34(11):4006–4021, March 2014.
- [3] Natsuki Yamanobe, Weiwei Wan, Ixchel G. Ramirez-Alpizar, Damien Petit, Tokuo Tsuji, Shuichi Akizuki, Manabu Hashimoto, Kazuyuki Nagata, and Kensuke Harada. A brief review of affordance in robotic manipulation research. *Advanced Robotics*, 31(19–20):1086–1101, October 2017.

- [4] Daniel Baldauf, He Cui, and Richard A. Andersen. The posterior parietal cortex encodes in parallel both goals for double-reach sequences. *Journal of Neuroscience*, 28(40):10081–10089, 2008.
- [5] David Meunier. Hierarchical modularity in human brain functional networks. *Frontiers in Neuroinformatics*, 3, 2009.
- [6] Claus C. Hilgetag and Alexandros Goulas. ‘hierarchy’ in the organization of brain networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1796):20190319, February 2020.
- [7] Nicholas P. Holmes and Charles Spence. The body schema and multisensory representation(s) of peripersonal space. *Cognitive Processing*, 5(2):94–105, June 2004.
- [8] Atsushi Yokoi and Jörn Diedrichsen. Neural organization of hierarchical motor sequence representations in the human neocortex. *Neuron*, 103(6):1178–1190.e7, September 2019.
- [9] Christine Assaiante, F. Barlaam, F. Cignetti, and M. Vaugoyeau. Body schema building during childhood and adolescence: A neurosensory approach. *Neurophysiologie Clinique = Clinical Neurophysiology*, 44(1):3–12, January 2014.
- [10] Atsushi Iriki, Michio Tanaka, and Yoshiaki Iwamura. Coding of modified body schema during tool use by macaque postcentral neurones. *NeuroReport*, 7(14):2325–2330, October 1996.
- [11] Shigeru Obayashi, Tetsuya Suhara, Koichi Kawabe, Takashi Okauchi, Jun Maeda, Yoshihide Akine, Hirotaka Onoe, and Atsushi Iriki. Functional brain mapping of monkey tool use. *NeuroImage*, 14(4):853–861, 2001.
- [12] Thomas A. Carlson, George Alvarez, Daw-an Wu, and Frans A.J. Verstraten. Rapid assimilation of external objects into the body schema. *Psychological Science*, 21(7):1000–1005, May 2010.
- [13] Lucilla Cardinali, Francesca Frassinetti, Claudio Brozzoli, Christian Urquizar, Alice C. Roy, and Alessandro Farnè. Tool-use induces morphological updating of the body schema. *Current Biology*, 19(13):478, 2009.
- [14] Giovanni Pezzulo, Francesco Donnarumma, Domenico Maisto, and Ivilin Stoianov. Planning at decision time and in the background during spatial navigation. *Current Opinion in Behavioral Sciences*, 29:69–76, 2019.
- [15] A David Redish. Vicarious trial and error. *Nature Reviews Neuroscience*, 17:147–159, 2016.
- [16] I Stoianov, C Pennartz, C Lansink, and G Pezzulo. Model-based spatial navigation in the hippocampus-ventral striatum circuit: a computational analysis. *Plos Computational Biology*, 14(9):1–28, 2018.
- [17] Ivilin Stoianov, Domenico Maisto, and Giovanni Pezzulo. The hippocampal formation as a hierarchical generative model supporting generative replay and continual learning. *Progress in Neurobiology*, 217:1–20, 2022.
- [18] Rajesh P.N. Rao and Dana H. Ballard. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87, 1999.
- [19] Jakob Hohwy. *The Predictive Mind*. Oxford University Press UK, 2013.
- [20] Andy Clark. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press, 01 2016.
- [21] Karl Friston and Stefan Kiebel. Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521):1211–1221, 2009.
- [22] Jakob Hohwy. New directions in predictive processing. *Mind and Language*, 35(2):209–223, 2020.
- [23] Andy Clark. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204, 2013.
- [24] Stewart Shipp. Neural elements for predictive coding. *Frontiers in Psychology*, 7, November 2016.
- [25] Beren Millidge, Anil Seth, and Christopher L Buckley. Predictive coding: a theoretical and experimental review, 2022.
- [26] Alexander Ororbia and Daniel Kifer. The neural coding framework for learning generative models. *Nature Communications*, 13(1), 2022.
- [27] Tommaso Salvatori, Ankur Mali, Christopher L. Buckley, Thomas Lukasiewicz, Rajesh P. N. Rao, Karl Friston, and Alexander Ororbia. Brain-inspired computational intelligence via predictive coding, 2023.
- [28] James C R Whittington and Rafal Bogacz. An approximation of the error backpropagation algorithm in a predictive coding network with local hebbian synaptic plasticity. *Neural Comput*, 29(5):1229–1262, March 2017.

- [29] James C.R. Whittington and Rafal Bogacz. Theories of Error Back-Propagation in the Brain. *Trends in Cognitive Sciences*, 23(3):235–250, 2019.
- [30] Beren Millidge, Alexander Tschantz, and Christopher L. Buckley. Predictive Coding Approximates Backprop Along Arbitrary Computation Graphs. *Neural Computation*, 34(6):1329–1368, 2022.
- [31] Jakob Hohwy, Andreas Roepstorff, and Karl Friston. Predictive coding explains binocular rivalry: An epistemological review. *Cognition*, 108(3):687–701, September 2008.
- [32] Karl J. Friston, Jean Daunizeau, James Kilner, and Stefan J. Kiebel. Action and behavior: A free-energy formulation. *Biological Cybernetics*, 102(3):227–260, 2010.
- [33] Karl Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.
- [34] Christopher L. Buckley, Chang Sub Kim, Simon McGregor, and Anil K. Seth. The free energy principle for action and perception: A mathematical review. *Journal of Mathematical Psychology*, 81:55–79, 2017.
- [35] Thomas Parr, Giovanni Pezzulo, and Karl J Friston. *Active inference: the free energy principle in mind, brain, and behavior*. Cambridge, MA: MIT Press, 2021.
- [36] Karl J. Friston, Jean Daunizeau, and Stefan J. Kiebel. Reinforcement learning or active inference? *PLoS ONE*, 4(7), 2009.
- [37] Karl Friston. What is optimal about motor control? *Neuron*, 72(3):488–498, 2011.
- [38] Rick A. Adams, Stewart Shipp, and Karl J. Friston. Predictions not commands: Active inference in the motor system. *Brain Structure and Function*, 218(3):611–643, 2013.
- [39] Harriet Brown, Karl Friston, and Sven Bestmann. Active inference, attention, and motor preparation. *Frontiers in Psychology*, 2(SEP):1–10, 2011.
- [40] Matteo Priorelli, Federico Maggiore, Antonella Maselli, Francesco Donnarumma, Domenico Maisto, Francesco Mannella, Ivilin Peev Stoianov, and Giovanni Pezzulo. Modeling motor control in continuous-time Active Inference: a survey. *IEEE Transactions on Cognitive and Developmental Systems*, pages 1–15, 2023.
- [41] Pablo Lanillos, Cristian Meo, Corrado Pezzato, Ajith Anil Meera, Mohamed Baioumy, Wataru Ohata, Alexander Tschantz, Beren Millidge, Martijn Wisse, Christopher L. Buckley, and Jun Tani. Active inference in robotics and artificial agents: Survey and challenges. *CoRR*, abs/2112.01871, 2021.
- [42] Tadahiro Taniguchi, Shingo Murata, Masahiro Suzuki, Dimitri Ognibene, Pablo Lanillos, Emre Ugur, Lorenzo Jamone, Tomoaki Nakamura, Alejandra Ciria, Bruno Lara, and Giovanni Pezzulo. World models and predictive coding for cognitive and developmental robotics: frontiers and challenges. *Advanced Robotics*, 37(13):780–806, June 2023.
- [43] Corrado Pezzato, Riccardo Ferrari, and Carlos Hernández Corbato. A novel adaptive controller for robot manipulators based on active inference. *IEEE Robotics and Automation Letters*, 5(2):2973–2980, 2020.
- [44] A. Maselli, P. Lanillos, and G. Pezzulo. Active inference unifies intentional and conflict-resolution imperatives of motor control. *PLOS Comput. Biol.*, 18(6), 2022.
- [45] Francesco Mannella, Federico Maggiore, Manuel Baltieri, and Giovanni Pezzulo. Active inference through whiskers. *Neural Networks*, 144:428–437, 2021.
- [46] Ajith Anil Meera, Filip Novicky, Thomas Parr, Karl Friston, Pablo Lanillos, and Noor Sajid. Reclaiming saliency: Rhythmic precision-modulated action and perception. *Frontiers in Neurorobotics*, 16:1–23, 2022.
- [47] Raphael Kaplan and Karl J. Friston. Planning and navigation as active inference. *Biological Cybernetics*, 112(4):323–343, 2018.
- [48] Rick A. Adams, Klaas Enno Stephan, Harriet R. Brown, Christopher D. Frith, and Karl J. Friston. The computational anatomy of psychosis. *Frontiers in Psychiatry*, 4, 2013.
- [49] Karl J. Friston, Thomas Parr, and Bert de Vries. The graphical brain: Belief propagation and active inference. 1(4):381–414, 2017.
- [50] Riccardo Proietti, Giovanni Pezzulo, and Alessia Tessari. An active inference model of hierarchical action understanding, learning and imitation. *Physics of Life Reviews*, 46:92–118, September 2023.

- [51] Francesco Donnarumma, Marcello Costantini, Ettore Ambrosini, Karl Friston, and Giovanni Pezzulo. Action perception as hypothesis testing. *Cortex*, 89:45–60, April 2017.
- [52] Giovanni Pezzulo, Leo D’Amato, Francesco Mannella, Matteo Priorelli, Toon Van de Maele, Ivilin Peev Stoianov, and Karl Friston. Neural representation in active inference: using generative models to interact with – and understand – the lived world. *Annals of the New York Academy of Sciences*, in press 2024.
- [53] Pablo Lanillos, Jordi Pages, and Gordon Cheng. Robot self/other distinction: active inference meets neural networks learning in a mirror. (Ecai), 2020.
- [54] Marc Toussaint and Amos Storkey. Probabilistic inference for solving discrete and continuous state Markov Decision Processes. *ACM International Conference Proceeding Series*, 148:945–952, 2006.
- [55] Marc Toussaint. Probabilistic inference as a model of planned behavior. *Künstliche Intelligenz*, 3/09:23–29, 2009.
- [56] Matthew Botvinick and Marc Toussaint. Planning as inference. *Trends in Cognitive Sciences*, 16(10):485–488, 2012.
- [57] Karl Friston. Hierarchical models in the brain. *PLoS Computational Biology*, 4(11), 2008.
- [58] Karl Friston, Klaas Stephan, Baojuan Li, and Jean Daunizeau. Generalised filtering. *Mathematical Problems in Engineering*, 2010:Article ID 621670, 34 p.–Article ID 621670, 34 p., 2010.
- [59] Stefano Ferraro, Toon Van de Maele, Pietro Mazzaglia, Tim Verbelen, and Bart Dhoedt. Disentangling shape and pose for object-centric deep active inference models, 2022.
- [60] Ruben S. van Bergen and Pablo L. Lanillos. Object-based active inference, 2022.
- [61] Toon Van de Maele, Tim Verbelen, Ozan undefinedatal, and Bart Dhoedt. Embodied object representation learning and recognition. *Frontiers in Neurorobotics*, 16, April 2022.
- [62] Toon Van de Maele, Tim Verbelen, Pietro Mazzaglia, Stefano Ferraro, and Bart Dhoedt. Object-centric scene representations using active inference, 2023.
- [63] Rick A. Adams, Eduardo Aponte, Louise Marshall, and Karl J. Friston. Active inference and oculomotor pursuit: The dynamic causal modelling of eye movements. *Journal of Neuroscience Methods*, 242:1–14, 2015.
- [64] Matteo Priorelli and Ivilin Peev Stoianov. Flexible Intentions: An Active Inference Theory. *Frontiers in Computational Neuroscience*, 17:1 – 41, 2023.
- [65] M. Priorelli and I.P. Stoianov. Dynamic inference by model reduction. *bioRxiv*, 2023.
- [66] Matteo Priorelli and Ivilin Peev Stoianov. Slow but flexible or fast but rigid? discrete and continuous processes compared. *Heliyon*, page e39129, October 2024.
- [67] Karl J. Friston, Richard Rosch, Thomas Parr, Cathy Price, and Howard Bowman. Deep temporal models and active inference. *Neuroscience and Biobehavioral Reviews*, 77(November 2016):388–402, 2017.
- [68] David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, April 2017.
- [69] Maxwell J. D. Ramstead, Dalton A. R. Sakthivadivel, Conor Heins, Magnus Koudahl, Beren Millidge, Lancelot Da Costa, Brennan Klein, and Karl J. Friston. On bayesian mechanics: a physics of and by beliefs. *Interface Focus*, 13(3), April 2023.
- [70] Cansu Sancaktar, Marcel A. J. van Gerven, and Pablo Lanillos. 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)*, pages 1–8, 2020.
- [71] Guillermo Oliver, Pablo Lanillos, and Gordon Cheng. An empirical study of active inference on a humanoid robot. *IEEE Transactions on Cognitive and Developmental Systems*, 8920(c):1–10, 2021.
- [72] Cristian Meo and Pablo Lanillos. Multimodal VAE active inference controller. *CoRR*, abs/2103.04412, 2021.
- [73] Thomas Rood, Marcel van Gerven, and Pablo Lanillos. A deep active inference model of the rubber-hand illusion. 2020.
- [74] Mohamed Baioumy, Paul Duckworth, Bruno Lacerda, and Nick Hawes. Active inference for integrated state-estimation, control, and learning. *arXiv*, 2020.

- [75] Cristian Meo, Giovanni Franzese, Corrado Pezzato, Max Spahn, and Pablo Lanillos. Adaptation through prediction: Multisensory active inference torque control. *IEEE Transactions on Cognitive and Developmental Systems*, 15(1):32–41, 2023.
- [76] Fred Bos, Ajith Anil Meera, Dennis Benders, and Martijn Wisse. Free Energy Principle for State and Input Estimation of a Quadcopter Flying in Wind. *Proceedings - IEEE International Conference on Robotics and Automation*, pages 5389–5395, 2022.
- [77] Ajith Anil Meera and Martijn Wisse. Dynamic expectation maximization algorithm for estimation of linear systems with colored noise. *Entropy*, 23(10), 2021.
- [78] Thomas Parr and Karl J. Friston. Active inference and the anatomy of oculomotion. *Neuropsychologia*, 111(January):334–343, 2018.
- [79] Léo Pio-Lopez, Ange Nizard, Karl Friston, and Giovanni Pezzulo. Active inference and robot control: A case study. *Journal of the Royal Society Interface*, 13(122), 2016.
- [80] Matteo Priorelli, Giovanni Pezzulo, and Ivilin Peev Stoianov. Deep kinematic inference affords efficient and scalable control of bodily movements. *Proceedings of the National Academy of Sciences of the United States of America*, 120, 2023.
- [81] Karl J Friston, Jérémie Mattout, and James Kilner. Action understanding and active inference. *Biological cybernetics*, 104(1-2):137–60, feb 2011.
- [82] Giovanni Pezzulo, Francesco Rigoli, and Karl J. Friston. Hierarchical Active Inference: A Theory of Motivated Control. *Trends in Cognitive Sciences*, 22(4):294–306, 2018.
- [83] M. Priorelli, G. Pezzulo, and I.P. Stoianov. Active vision in binocular depth estimation: A top-down perspective. *Biomimetics*, 8(5), 2023.
- [84] Emanuel Todorov. Optimality principles in sensorimotor control. *Nature Neuroscience*, 7:907–915, 2004.
- [85] Mareike Floegel, Johannes Kasper, Pascal Perrier, and Christian A. Kell. How the conception of control influences our understanding of actions. *Nature Reviews Neuroscience*, 24(May):313–329, 2023.
- [86] Giuseppe Vallar, Elie Lobel, Gaspare Galati, Alain Berthoz, Luigi Pizzamiglio, and Denis Le Bihan. A fronto-parietal system for computing the egocentric spatial frame of reference in humans. *Experimental Brain Research*, 124(3):281–286, January 1999.
- [87] James R. Hinman, G. William Chapman, and Michael E. Hasselmo. Neuronal representation of environmental boundaries in egocentric coordinates. *Nature Communications*, 10(1), June 2019.
- [88] Karl J. Friston, Thomas Parr, Yan Yufik, Noor Sajid, Catherine J. Price, and Emma Holmes. Generative models, linguistic communication and active inference. *Neuroscience and Biobehavioral Reviews*, 118:42–64, November 2020.
- [89] Beren Millidge, Mahyar Osanlouy, and Rafal Bogacz. Predictive Coding Networks for Temporal Prediction. pages 1–59, 2023.
- [90] Matteo Priorelli and Ivilin Peev Stoianov. Deep hybrid models: infer and plan in the real world. *arXiv*, 2024.
- [91] Rajesh P. N. Rao, Dimitrios C. Gklezakos, and Vishwas Sathish. Active predictive coding: A unified neural framework for learning hierarchical world models for perception and planning, 2022.
- [92] Ares Fisher and Rajesh P N Rao. Recursive neural programs: A differentiable framework for learning compositional part-whole hierarchies and image grammars. *PNAS Nexus*, 2(11), October 2023.
- [93] Giacomo Rizzolatti and Laila Craighero. The mirror-neuron system. *Annu Rev Neurosci*, 27:169–192, 2004.
- [94] James M. Kilner, Karl J. Friston, and Chris D. Frith. Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, 8(3):159–166, April 2007.
- [95] Jeff Hawkins, Subutai Ahmad, and Yuwei Cui. A theory of how columns in the neocortex enable learning the structure of the world. *Frontiers in Neural Circuits*, 11, 2017.
- [96] Lancelot Da Costa, Thomas Parr, Noor Sajid, Sebastian Veselic, Victorita Neacsu, and Karl Friston. Active inference on discrete state-spaces: A synthesis. *Journal of Mathematical Psychology*, 99, 2020.

- [97] Ryan Smith, Karl J. Friston, and Christopher J. Whyte. A step-by-step tutorial on active inference and its application to empirical data. *Journal of Mathematical Psychology*, 107:102632, 2022.
- [98] Thomas Parr, Rajeet Vijay Rikhye, Michael M. Halassa, and Karl J. Friston. Prefrontal Computation as Active Inference. *Cerebral Cortex*, 30(2):682–695, 2020.
- [99] K. J. Friston, L. Harrison, and Will Penny. Dynamic causal modelling. *NeuroImage*, 19(4):1273–1302, 2003.
- [100] Karl Friston and Will Penny. Post hoc Bayesian model selection. *NeuroImage*, 56(4):2089–2099, 2011.
- [101] Karl Friston, Thomas Parr, and Peter Zeidman. Bayesian model reduction. pages 1–32, 2018.
- [102] M.J. Rosa, K. Friston, and W. Penny. Post-hoc selection of dynamic causal models. *Journal of Neuroscience Methods*, 208(1):66–78, June 2012.
- [103] Thomas Parr and Karl J. Friston. The Discrete and Continuous Brain: From Decisions to Movement—And Back Again Thomas. *Neural Computation*, 30:2319–2347, 2018.
- [104] T. Parr and K. J. Friston. The computational pharmacology of oculomotion. *Psychopharmacology (Berl.)*, 236(8):2473–2484, August 2019.
- [105] A. Tschantz, L. Barca, D. Maisto, C. L. Buckley, A. K. Seth, and G. Pezzulo. Simulating homeostatic, allostatic and goal-directed forms of interoceptive control using active inference. *Biological Psychology*, 169:108266, 2022.
- [106] Ozan Çatal, Tim Verbelen, Toon Van de Maele, Bart Dhoedt, and Adam Safron. Robot navigation as hierarchical active inference. *Neural Networks*, 142:192–204, 2021.
- [107] Sheida Nozari, Ali Krayani, Pablo Marin-Plaza, Lucio Marcenaro, David Martin Gomez, and Carlo Regazzoni. Active inference integrated with imitation learning for autonomous driving. *IEEE Access*, 10:49738–49756, 2022.
- [108] Sheida Nozari, Ali Krayani, Pablo Marin, Lucio Marcenaro, David Martin Gomez, and Carlo Regazzoni. Exploring action-oriented models via active inference for autonomous vehicles. July 2023.
- [109] Ali Krayani, Khalid Khan, Lucio Marcenaro, Mario Marchese, and Carlo Regazzoni. A goal-directed trajectory planning using active inference in uav-assisted wireless networks. *Sensors*, 23(15):6873, August 2023.
- [110] Ali Krayani, Khalid Khan, Lucio Marcenaro, Mario Marchese, and Carlo Regazzoni. Self-supervised path planning in uav-aided wireless networks based on active inference. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, volume 10, page 13181–13185. IEEE, April 2024.
- [111] Felix Obite, Ali Krayani, Atm S. Alam, Lucio Marcenaro, Arumugam Nallanathan, and Carlo Regazzoni. Intelligent resource allocation for uav-based cognitive noma networks: An active inference approach. In *2023 IEEE Future Networks World Forum (FNWF)*, volume 182, page 1–7. IEEE, November 2023.
- [112] Ali Krayani, Atm S. Alam, Lucio Marcenaro, Arumugam Nallanathan, and Carlo Regazzoni. A novel resource allocation for anti-jamming in cognitive-uavs: An active inference approach. *IEEE Communications Letters*, 26(10):2272–2276, October 2022.
- [113] Corrado Pezzato, Carlos Hernández Corbato, Stefan Bonhof, and Martijn Wisse. Active inference and behavior trees for reactive action planning and execution in robotics. *IEEE Transactions on Robotics*, 39(2):1050–1069, April 2023.
- [114] Poppy Collis, Ryan Singh, Paul F Kinghorn, and Christopher L Buckley. Learning in hybrid active inference models, 2024.
- [115] Nathan F. Lepora and Giovanni Pezzulo. Embodied choice: How action influences perceptual decision making. *PLOS Computational Biology*, 11(4):e1004110, April 2015.
- [116] Matteo Priorelli, Ivilin Peev Stoianov, and Giovanni Pezzulo. Embodied decisions as active inference. June 2024.
- [117] Takuya Isomura, Thomas Parr, and Karl Friston. Bayesian filtering with multiple internal models: Toward a theory of social intelligence. *Neural Computation*, 31(12):2390–2431, 2019.
- [118] Harriet Feldman and Karl J. Friston. Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4, 2010.

- [119] Thomas Parr, David A. Benrimoh, Peter Vincent, and Karl J. Friston. Precision and false perceptual inference. *Frontiers in Integrative Neuroscience*, 12, September 2018.
- [120] F. Fattapposta, G. Amabile, M. V. Cordischi, D. Di Venanzio, A. Foti, F. Pierelli, C. D’Alessio, F. Pigozzi, A. Parisi, and C. Morrocutti. Long-term practice effects on a new skilled motor learning: An electrophysiological study. *Electroencephalography and Clinical Neurophysiology*, 99(6):495–507, 1996.
- [121] Francesco Di Russo, Sabrina Pitzalis, Teresa Aprile, and Donatella Spinelli. Effect of practice on brain activity: An investigation in top-level rifle shooters. *Medicine and Science in Sports and Exercise*, 37(9):1586–1593, 2005.
- [122] Ann M. Graybiel. Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, 31:359–387, 2008.
- [123] Karl J. Friston, Thomas Parr, Conor Heins, Axel Constant, Daniel Friedman, Takuya Isomura, Chris Fields, Tim Verbelen, Maxwell Ramstead, John Clippinger, and Christopher D. Frith. Federated inference and belief sharing. *Neuroscience & Biobehavioral Reviews*, 156:105500, 2024.
- [124] Toon Van de Maele, Bart Dhoedt, Tim Verbelen, and Giovanni Pezzulo. Integrating cognitive map learning and active inference for planning in ambiguous environments. In *Active Inference*, pages 204–217, Cham, 2024. Springer Nature Switzerland.
- [125] Daria de Tinguy, Toon Van de Maele, Tim Verbelen, and Bart Dhoedt. Spatial and temporal hierarchy for autonomous navigation using active inference in minigrid environment. *Entropy*, 26(1):83, January 2024.
- [126] Karl Friston, Lancelot Da Costa, Danijar Hafner, Casper Hesp, and Thomas Parr. Sophisticated inference. *Neural Computation*, 33(3):713–763, 2021.
- [127] Matteo Priorelli, Ivilin Peev Stoianov, and Giovanni Pezzulo. Learning and embodied decisions in active inference. August 2024.
- [128] Daniele Caligiore, Tomassino Ferrauto, Domenico Parisi, Neri Accornero, Marco Capozza, and Gianluca Baldassarre. Using motor babbling and hebb rules for modeling the development of reaching with obstacles and grasping. 2008.
- [129] Matteo Priorelli and Ivilin Peev Stoianov. Efficient motor learning through action-perception cycles in deep kinematic inference. In *Active Inference*, pages 59–70. Springer Nature Switzerland, 2024.
- [130] Domenico Maisto, Francesco Donnarumma, and Giovanni Pezzulo. Interactive inference: A multi-agent model of cooperative joint actions. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 54(2):704–715, 2024.
- [131] Aswin Paul, Noor Sajid, Lancelot Da Costa, and Adeel Razi. On efficient computation in active inference. *Expert Systems with Applications*, 253:124315, November 2024.
- [132] Linxing Preston Jiang and Rajesh P. N. Rao. Dynamic predictive coding: A model of hierarchical sequence learning and prediction in the neocortex. *bioRxiv*, 2023.
- [133] Alexander Ororbia and Ankur Mali. Active Predicting Coding: Brain-Inspired Reinforcement Learning for Sparse Reward Robotic Control Problems. 2022.
- [134] Beren Millidge. Combining Active Inference and Hierarchical Predictive Coding: a Tutorial Introduction and Case Study. *PsyArXiv*, 2019.
- [135] Kai Ueltzhöffer. Deep Active Inference. pages 1–40, 2017.
- [136] Ozan Çatal, Johannes Nauta, Tim Verbelen, Pieter Simoens, and B. Dhoedt. Bayesian policy selection using active inference. *ArXiv*, abs/1904.08149, 2019.
- [137] Stefano Ferraro, Toon Van de Maele, Tim Verbelen, and Bart Dhoedt. Symmetry and complexity in object-centric deep active inference models. *Interface Focus*, 13(3), April 2023.
- [138] Kai Yuan, Karl Friston, Zhibin Li, and Noor Sajid. Hierarchical generative modelling for autonomous robots. *Research Square*, 2023.
- [139] Beren Millidge. Deep active inference as variational policy gradients. *Journal of Mathematical Psychology*, 96, 2020.
- [140] Zafeirios Fountas, Noor Sajid, Pedro A.M. Mediano, and Karl Friston. Deep active inference agents using Monte-Carlo methods. *Advances in Neural Information Processing Systems*, 2020-Decem(NeurIPS), 2020.

- [141] Théophile Champion, Marek Grześ, Lisa Bonheme, and Howard Bowman. Deconstructing deep active inference. 2023.
- [142] Aleksey Zelenov and Vladimir Krylov. Deep active inference in control tasks. In *2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, pages 1–3, 2021.
- [143] Stephen Grossberg and Praveen K. Pilly. Temporal dynamics of decision-making during motion perception in the visual cortex. *Vision Research*, 48(12):1345–1373, June 2008.
- [144] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [145] Giovanni Pezzulo, Thomas Parr, Paul Cisek, Andy Clark, and Karl Friston. Generating meaning: active inference and the scope and limits of passive ai. *Trends in Cognitive Sciences*, 28(2):97–112, February 2024.
- [146] Karl J. Friston, Lancelot Da Costa, Alexander Tschantz, Alex Kiefer, Tommaso Salvatori, Victorita Neacsu, Magnus Koudahl, Conor Heins, Noor Sajid, Dimitrije Markovic, Thomas Parr, Tim Verbelen, and Christopher L Buckley. Supervised structure learning. 2023.
- [147] Adam N Sanborn, Thomas L Griffiths, and DJ Navarro. Rational approximations to rational models: alternative algorithms for category learning. *Psychological Review*, 117(4):1144–1167, 2010.
- [148] Ivilin Stoianov, Aldo Genovesio, and Giovanni Pezzulo. Prefrontal goal codes emerge as latent states in probabilistic value learning. *Journal of Cognitive Neuroscience*, 28(1):140–157, 2016.