DYNAMIC INFERENCE BY MODEL REDUCTION

Matteo Priorelli
Institute of Cognitive Sciences and Technologies
National Research Council of Italy
Padova, 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

How do we infer which one of several targets another agent is following? And how are we capable of grasping an object on the fly? Reducing a model as complex as the surrounding dynamic environment into a small set of simpler hypotheses is a reasonable cognitive solution, but how can the brain compare and choose among dynamic hypotheses? Recent advances in Bayesian Model Reduction have led to innovative solutions to actively infer the state of affairs of the world and perform discrete planning with continuous signals, but dealing with highly dynamic contexts is a difficult matter. We propose that choosing among flexible hypotheses is possible by using reduced priors sampled from the dynamics of a generative model. Each reduced prior corresponds to an alternative future world constantly generated from the current observations, which the agent can use to accumulate evidence for a discrete hypothesis. We tested the approach on two everyday tasks: inferring a dynamic trajectory and grasping a moving object, which the model solved with high accuracy. Our study may thus shed light on how agents can smoothly generate and infer static intentions that result in dynamic behaviors, and on the specific role of intention precisions in motor learning.

1 Introduction

How can biological organisms dynamically infer a representation of the world that allows high-level planning and decision-making? A recent theory, called Active Inference, proposes that in order to survive, all organisms are constantly involved in minimizing a quantity called *free energy* [1, 2, 3, 4]. In order to restrict the possible states that allow us to live, the nervous system is supposed to maintain an internal generative model about how hidden states generate observations and how such states evolve. The resulting predictions are compared with the observations to form a prediction error, and depending on how this error is minimized gives rise to either one or the other of the two essential components of control, i.e., perception and action [5, 6].

Dealing with a continuous environment is possible by expressing the hidden states in terms of generalized coordinates (e.g., position, velocity, acceleration, and so on), allowing to represent the environment dynamics with high fidelity [7, 8]. However, this dynamics can be only minimized through the free energy of present states and observations, hence the agent is not capable of any sort of decision-making. The latter is possible by augmenting a continuous model with a discretized representation of the environment: in this case, the high-level hidden states do not represent instantaneous trajectories, but discrete future states. Planning then arises when future observations are treated as hidden states – an approach called planning as inference [9, 10, 11] – and when the latter are minimized with the free energy that the agent expects to perceive in the future [12, 13, 14]. Using a discrete model for planning poses however the problem of how to generate continuous causes from discrete observations. This has been usually done through Bayesian Model Reduction (BMR) [15], a fundamental technique that has been utilized in several areas over the course of the years [16, 17, 18, 19, 20, 21, 22, 23]. This technique consists of reducing a complex distribution into a smaller and simpler set of hypotheses that have the same likelihood and differ only in their priors.

Models that combine discrete and continuous computations through BMR are called *hybrid*, and a few applications can be found in the literature simulating epistemic foraging in reading, [24, 25], multi-step reaching movements in neurological disorders [1], saccades and visual sampling [26, 27], active listening [28], and interoceptive control [29]. However, the reduced models that are used for comparison with the full complex distribution are usually determined before the inferential process. As a consequence, an agent can only infer a discrete cause among a fixed set of hypotheses; although it could be possible that such reduced models are adaptable, this process would be generally slow and could not deal efficiently with highly dynamic environments. Here, we propose an alternative solution that

combines the two frameworks at a different level leading to further implications and parallelisms. This approach can be used in several dynamics contexts, and we demonstrate its capabilities in two common and non-trivial tasks: inferring the trajectory of another agent that wants to follow one of many moving targets, and grasping a moving object.

2 Results

2.1 Dynamic Reduced Models

Figure 1a shows the factor graph of a conventional hybrid model. For simplicity, we do not consider the whole discrete model that also includes transitions among discrete hidden states; instead, we assume that the discrete states $\boldsymbol{s}$ are sampled from a categorical distribution:

$$p(\boldsymbol{s}) = Cat(\boldsymbol{D}) \tag{1}$$

where $\boldsymbol{D}$ encodes the initial state. The continuous model is supposed to learn a dynamic representation of the environment modeled by a system of stochastic differential equations:

$$\begin{aligned} \tilde{\boldsymbol{o}} &= \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}) + \tilde{\boldsymbol{w}}_o \\ \mathcal{D}\tilde{\boldsymbol{x}} &= \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}) + \tilde{\boldsymbol{w}}_x \end{aligned} \tag{2}$$

where $\tilde{\boldsymbol{o}}$, $\tilde{\boldsymbol{g}}$, $\tilde{\boldsymbol{x}}$, $\tilde{\boldsymbol{f}}$, $\tilde{\boldsymbol{v}}$ are respectively the observations, likelihood function, continuous hidden states, dynamics function, and hidden causes. Note that the symbol $\tilde{}$ denotes a generalized representation encoding instantaneous trajectories up to a certain order, i.e. $\tilde{\boldsymbol{x}} = [\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{x}'', \dots]$, and that $\mathcal{D}$ is a shift differential operator such that $\mathcal{D}\tilde{\boldsymbol{x}} = [\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}''' \dots]$. Further, $\tilde{\boldsymbol{w}}_o$ and $\tilde{\boldsymbol{w}}_x$ are noise terms sampled from Gaussian distributions.

The continuous probability distributions of the hybrid model are assumed to be the following:

$$\begin{aligned} p(\tilde{\boldsymbol{o}}|\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{\Sigma}}_o) \\ p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}}) &= \mathcal{N}(\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}), \tilde{\boldsymbol{\Sigma}}_x) \\ p(\tilde{\boldsymbol{v}}|\boldsymbol{s}) &= \mathcal{N}(\boldsymbol{\eta}, \tilde{\boldsymbol{\Sigma}}_v) \end{aligned} \tag{3}$$

The last distribution tells us that the hidden causes are generated from discrete hidden states $\boldsymbol{s}$ through a Gaussian distribution with mean $\boldsymbol{\eta}$; this is considered a full model over a complex set of hypotheses. We can infer the most likely hidden cause under the current discrete states by assuming that the agent maintains M reduced models:

$$p(\tilde{\boldsymbol{v}}|\boldsymbol{s}, m) = \mathcal{N}(\boldsymbol{\eta}_m, \tilde{\boldsymbol{\Sigma}}_m) \tag{4}$$

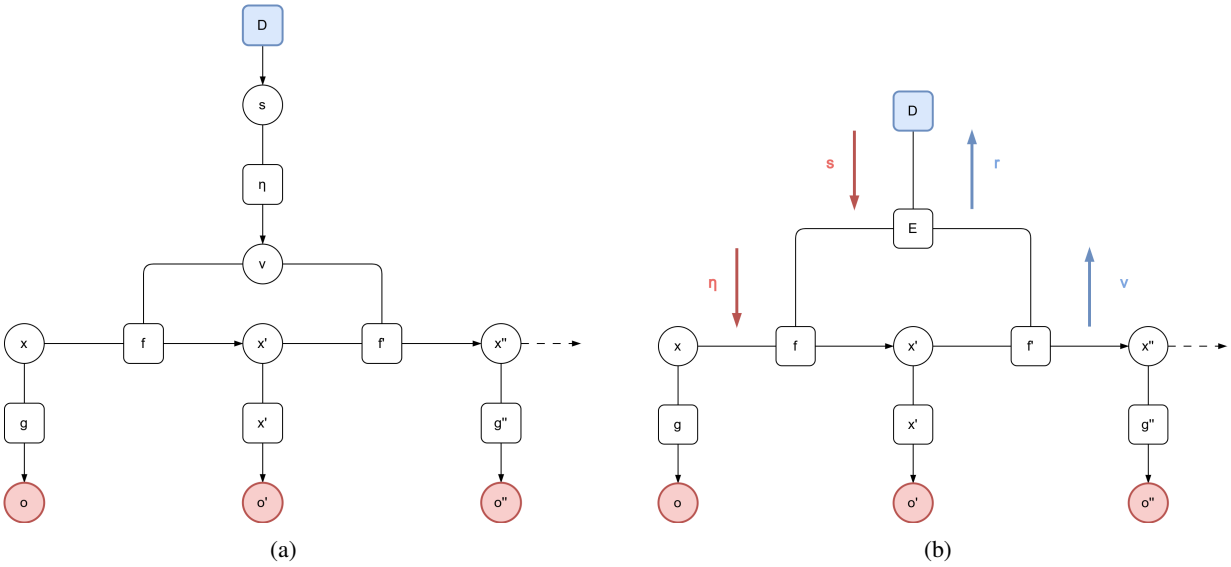
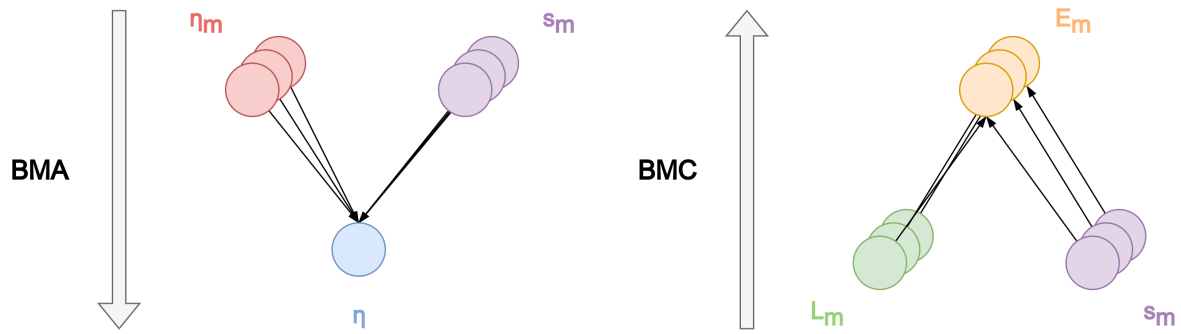


Figure 1: (a) Factor graph of a conventional hybrid model. For simplicity, only a prior categorical distribution is displayed. (b) A graph highlighting the top-down (red) and bottom-up (blue) messages of the BMR. The former ($\boldsymbol{\eta}$) compute an average between the reduced priors $\boldsymbol{\eta}_m$ (not displayed here) and the hidden states $\boldsymbol{s}$. The latter ($\boldsymbol{r}$) return a comparison between the reduced priors and the posterior over hidden causes $\boldsymbol{v}$ through the free energy $\boldsymbol{E}$.



(a) (b)

Figure 2: (a) The Bayesian Model Average (left) combines priors $\boldsymbol{\eta}_m$ of reduced models with probabilities s_m of a categorical distribution. (b) The Bayesian Model Comparison (right) compares the prior expectation of each model encoded in the surprise $-\ln s_m$ with the respective log evidence $\mathcal{L}_m$.

with the same likelihood of the full model but differing in their prior. As highlighted in Figure 1b, the exchange of top-down and bottom-up messages between the discrete and continuous representations follows the derivations of the BMR technique (see the Methods section for more details): after having defined appropriate approximate posteriors, the former are computed through a Bayesian Model Average (BMA) between the reduced priors and the discrete state models:

$$\boldsymbol{\eta} = \sum_m s_m \boldsymbol{\eta}_m \tag{5}$$

Instead, the latter are found through a Bayesian Model Comparison (BMC) between the reduced priors and the posterior over the hidden causes:

$$E_m = -\ln s_m - \int_0^T \mathcal{L}_m dt \qquad\qquad \boldsymbol{r} = \sigma(-\boldsymbol{E}) \tag{6}$$

where E_m and $\mathcal{L}_m$ are respectively the free energy and log evidence of the m-th reduced model, and $\boldsymbol{r}$ is the bottom-up message passed through a softmax function σ to get a proper probability.

What Equation 4 represents is that the agent has some prior hypotheses or assumptions over the possible causes that could have generated what it is currently perceiving. It hence makes its best guess for the lower levels by combining such continuous hypotheses and weighting them with the respective discrete probabilities (see Equation 5 and Figure 2a). The other side of the flow works in a specular way: the real assumption that best explains the current situation is found by first maintaining, for a certain amount of time, a score – i.e., the log evidence $\mathcal{L}_m$ – of the true posterior of the hidden causes with respect to the set of hypotheses which the agent can choose from. Then, it compares the score with its prior expectations encoded in the surprise $-\ln s_m$ (see Figure 2b).

As always, a delicate balance exists between expectations and evidences. In addition to this, a major problem is that the agent's set of hypotheses is determined a-priori, so that an observation can be explained only with an average of fixed continuous states. Considering the two examples mentioned in the Introduction, let us assume that the goal of an agent is to randomly choose at each trial one of several targets and reach it, while the goal of a second agent is to infer the trajectory of the other depending on the same set of hypotheses. The problem can be easily addressed by considering the target positions as reduced priors and comparing them with a hidden cause inferred through the observations. However, the agent would fail the inference if the targets are constantly moving, because it cannot rapidly update its reduced models. Similar behavior occurs in the second example: in this case, a discrete model is needed in order to infer the correct sequence of actions needed to grasp the object, i.e., reach and then grasp. Again, the task is easy if the object is not moving during the first action; in the latter condition however, the agent would assign the old object's position to the same discrete state, thus failing the task. Further, the object might move not only during the first reaching movement but also during the grasping phase, so the agent has to carefully balance both actions depending on the object's position at each continuous instant. The accumulation period of Equation 6 actually poses a limit on how frequently the model can either change the full priors that have to be realized, or compare the constantly changing evidence with its fixed set of reduced hypotheses.

In short, a comparison between hidden causes and reduced models can only deal with static conditions, and updating the reduced continuous signals that the agent uses to score its discrete probabilities is needed if the environment contains dynamic elements. How then to perform planning and inference with dynamic representations?

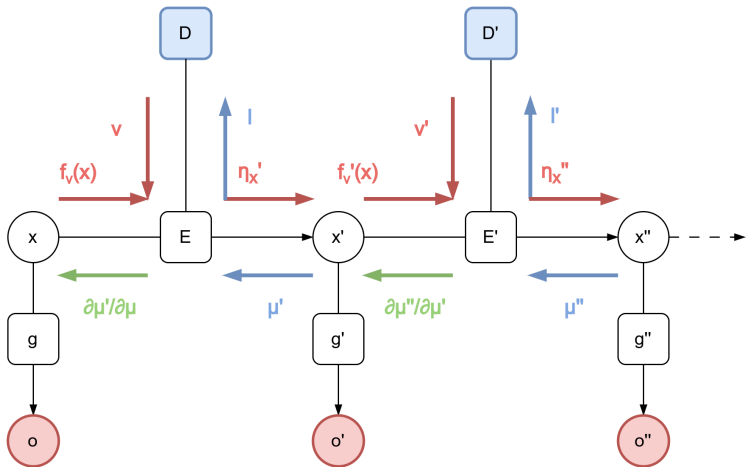


Figure 3: Graph highlighting the exchange of messages between dynamic reduced models. Here, the hidden causes $\boldsymbol{v}$ encode a categorical distribution and the reduced priors are continuously generated from the dynamics functions of the hidden states $\boldsymbol{f}_v(\boldsymbol{x})$. The full prior $\boldsymbol{\eta}'_x$ represents an average velocity that depends on the current observations and the agent's intentions. Instead, the gradient $\partial_\mu \boldsymbol{\mu}'$ indicates how this velocity is mapped back into its previous order $\boldsymbol{\mu}$. Last, the bottom-up message $\boldsymbol{l}$ conveys the most likely agent's intention that could have generated the current velocity.

First, we may consider the hidden causes $\boldsymbol{v}$ as directly generated from a categorical distribution:

$$p(\tilde{\boldsymbol{v}}) = Cat(\tilde{\boldsymbol{D}}) \tag{7}$$

At this point, we define V reduced prior probability distributions and a full prior model concerning the hidden states:

$$\begin{aligned} p(\tilde{\boldsymbol{x}}|v) &= \mathcal{N}(\tilde{\boldsymbol{f}}_v(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{\Sigma}}_{x,v}) \\ p(\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{\eta}}, \tilde{\boldsymbol{\Sigma}}_x) \end{aligned} \tag{8}$$

where $\tilde{\boldsymbol{\eta}}$ is the full prior and $\tilde{\boldsymbol{f}}_v(\tilde{\boldsymbol{x}})$ is the dynamics function of the v-th reduced model. Note that the latter is not directly conditioned over the hidden causes, but each reduced prior is associated with a specific cause. Then, we define the corresponding posterior models:

$$\begin{aligned} \tilde{q}(\tilde{\boldsymbol{x}}|v) &= \mathcal{N}(\tilde{\boldsymbol{\mu}}_v, \tilde{\boldsymbol{C}}_{x,v}) \\ \tilde{q}(\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{C}}_x) \end{aligned} \tag{9}$$

Note that $\tilde{\boldsymbol{\mu}}_v$ is not the posterior over the hidden causes – which is denoted by $\tilde{\boldsymbol{\nu}}$ – but the posterior over the hidden states of the v-th dynamics function. What Equations 8 and 9 imply can be better understood from Figure 3, which highlights three different kinds of message exchange between discrete and continuous representations.

During the forward pass, the discrete hidden causes for each temporal order are used to perform a BMA with as many dynamically generated trajectories as their models:

$$\tilde{\boldsymbol{\eta}} = \sum_v \tilde{\boldsymbol{v}}_v \tilde{\boldsymbol{f}}_v(\tilde{\boldsymbol{\mu}}) \tag{10}$$

The consequence is that, although a discrete model encodes the same state as before, the latter is associated with a different trajectory of the hidden states at each continuous step. In other words, the agent does not maintain fixed priors over environmental causes, but dynamic priors that define possible future trajectories based on the one that it is currently perceiving. In this way, a discrete model can impose high-level actions that result in a dynamic behavior for the whole continuous period of evidence accumulation.

As concerns the latter, each model of the hidden causes is scored as before, but now the posterior over the current trajectory is compared with the ones generated by the dynamics functions of the hidden states:

$$\tilde{\boldsymbol{\mathcal{L}}}_v = \frac{1}{2}(\tilde{\boldsymbol{\mu}}_v^T \tilde{\boldsymbol{P}}_{x,v} \tilde{\boldsymbol{\mu}}_v - \tilde{\boldsymbol{f}}_v(\tilde{\boldsymbol{\mu}})^T \tilde{\boldsymbol{\Pi}}_{x,v} \tilde{\boldsymbol{f}}_v(\tilde{\boldsymbol{\mu}}) - \tilde{\boldsymbol{\mu}}^T \tilde{\boldsymbol{P}}_x \tilde{\boldsymbol{\mu}} + \tilde{\boldsymbol{\eta}}^T \tilde{\boldsymbol{\Pi}}_x \tilde{\boldsymbol{\eta}}) \tag{11}$$

expressed in terms of precision $\boldsymbol{P}$ and $\boldsymbol{\Pi}$. At this point, for each temporal order a BMC is performed between the prior surprise – in this case, the elements of $\tilde{\boldsymbol{D}}$ – and the log evidences accumulated over a time T:

$$\begin{aligned} \tilde{\boldsymbol{l}}_v &= \int_0^T \tilde{\boldsymbol{\mathcal{L}}}_v dt \\ \tilde{\boldsymbol{v}} &= \sigma(-\tilde{\boldsymbol{E}}) = \sigma(\ln \tilde{\boldsymbol{D}} + \tilde{\boldsymbol{l}}) \end{aligned} \tag{12}$$

The bottom-up message $\tilde{\boldsymbol{l}}$ conveys the most likely cause $\tilde{\boldsymbol{v}}_v$ that could have produced the current trajectory. But since the latter continuously generates the reduced priors, a cause can be inferred which corresponds to a dynamic pathway for the whole period of evidence accumulation.

About the other backward pass, the dynamics prediction error between the actual trajectory and the full prior:

$$\tilde{\boldsymbol{\varepsilon}}_x = \mathcal{D}\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{\eta}} \tag{13}$$

backpropagates through the gradient of the full prior with respect to the posterior over the hidden states:

$$\dot{\tilde{\boldsymbol{\mu}}} \propto \frac{\partial \tilde{\boldsymbol{\eta}}}{\partial \tilde{\boldsymbol{\mu}}}^T \tilde{\boldsymbol{\Pi}}_x \tilde{\boldsymbol{\varepsilon}}_x \tag{14}$$

This gradient encodes how an average future trajectory is mapped back to the current one. Specifically, for each temporal order the most likely continuous state that could have been generated from the previous order is inferred.

There are two specular consequences of this architecture, related to the two previous examples. The intention of another agent whose goal is to reach one of several moving targets can be found by generating and comparing dynamic trajectories depending on the positions of the other's hand and every target, both continuously inferred from sensory observations. Similarly, grasping a moving object unfolds into the following steps: (i) a discrete model predicts a hidden cause corresponding to the reaching action; (ii) the hidden cause steers the lower-level dynamics by generating an average trajectory based on the current one; (iii) the biased trajectory backpropagates in terms of dynamics prediction errors, eventually generating proprioceptive predictions needed for movement; (iv) after a period T, the discrete backward pass infers the most likely hidden causes based on the accumulated evidence – i.e., it signals whether the cause has been realized – and the discrete model finally predicts the next (grasping) action.

Further, a parallelism arises with the discrete framework: if v^v is set to 1 and all the others to 0, the hidden states will be subject only to cause v; and if multiple hidden causes are active, the hidden states will be pulled toward a combination of them, in the same way that a high-level policy generates the transitions for the discrete hidden states below. In other words, we can consider the hidden causes as encoding both the agent's intentions and its assumptions. Notably, the precisions $\tilde{\boldsymbol{\Pi}}_{x,v}$ of Equation 11 can be interpreted with a perspective specular to the sensory precisions $\boldsymbol{\Pi}_o$. If a dynamics function has a strong precision, it means that it is a good option for minimizing the free energy in the current context or, more clearly, for both inferring the current agent's trajectory and for realizing its desired goal. Conversely, a low precision means that the agent is not confident about using that trajectory for the task considered – e.g., a grasping action when the object is out of reach – or that it cannot be helpful to understand the state of affairs of the world – e.g., a target that is far away from the hand of another agent.

In summary, two concurrent processes exist: a fast process that infers and imposes a trajectory depending on high-level expectations and low-level evidences; and a slow process that scores each trajectory and learns the most appropriate one for a given context.

2.2 Two examples of Dynamic Inference

Here, we briefly describe two examples of dynamic inference. In the first one, we infer the intention of an agent based on the trajectories of its arm and two targets. The second example describes the control process of an agent aiming to reach and grasp a moving object.

2.2.1 Trajectory Inference

We first consider a situation in which an agent is reaching one of two moving targets and we want to infer which one it is following. The dynamic nature of the task can be tackled through the definition of intentions [30, 31], defined as functions that generate a possible future state from the current continuous belief. In Active Inference, beliefs can be updated not only by inferring the causes of sensory signals, but also by defining priors about their evolution over time. The dynamics of a hidden state does not necessarily need to be the same as the true generative process: indeed, it is this discrepancy that makes the agent able to fulfill prior expectations. The belief dynamics can be seen as a combination

of two components (although this distinction is purely formal as from the agent's point of view there is only a single dynamics): (i) an evolution of the hidden states learned from past experiences and based on plausible assumptions over the environment – e.g., if I see a ball moving, I may have a prior that it will follow a linear trajectory; (ii) a distorted vision of this dynamics that depends on the agent's desires – e.g., I can think that my hand is pulled toward the ball, and even if such a force does not exist in the external world, it will drive my belief toward my expectations, ultimately affecting the movements of my hand.

Figure 4 shows a graphical representation of a model that can perform the presented task. This is a basic task that does not involve high-level planning, thus we did not include a discrete model and we considered an uninformative prior over the hidden causes. The agent consists of an arm with a single Degree of Freedom (DoF). Its sensory modalities are: (i) a proprioceptive observation o_p for the arm's joint angle; (ii) a visual observation $\boldsymbol{o}_v$ encoding the Cartesian positions of the hand and two targets.

Since extrinsic beliefs are not strictly necessary for this task, we express a belief over hidden states $\boldsymbol{\mu}$ in intrinsic coordinates (e.g., joint angles) only, which thus generate and integrate both sensory modalities. This belief encodes information about all the three elements of the environment:

$$\boldsymbol{\mu} = [\mu_a \quad \mu_{t1} \quad \mu_{t2}] \tag{15}$$

Proprioceptive predictions – needed to infer the kinematic configuration of the body's agent – are produced through a function g_p that simply extracts the arm's joint angle from the belief:

$$p_p = g_p(\boldsymbol{\mu}) = \mu_a + w_p \tag{16}$$

Instead, visual predictions are generated through forward kinematics $\boldsymbol{J}$ of the hidden states components:

$$\begin{aligned} \boldsymbol{p}_v = \boldsymbol{g}_v(\boldsymbol{\mu}) &= [\boldsymbol{J}(\mu_a) \quad \boldsymbol{J}(\mu_{t1}) \quad \boldsymbol{J}(\mu_{t2})] + \boldsymbol{w}_v \\ \boldsymbol{J}(\theta) &= [lc_\theta \quad ls_\theta] \end{aligned} \tag{17}$$

where l is the arm length, while s_θ and c_θ are the sine and cosine of the angle. In other words, the visual likelihood $\boldsymbol{g}_v$ performs, for each element, a forward kinematics from joint angles to Cartesian positions. While the interpretation of the first component of the hidden states (i.e., the agent's body) is intuitive, the other components can be seen as encoding a possible inferred configuration that could have generated the object's position. In short, they may be considered as states maintaining the agent's affordances for a particular object.

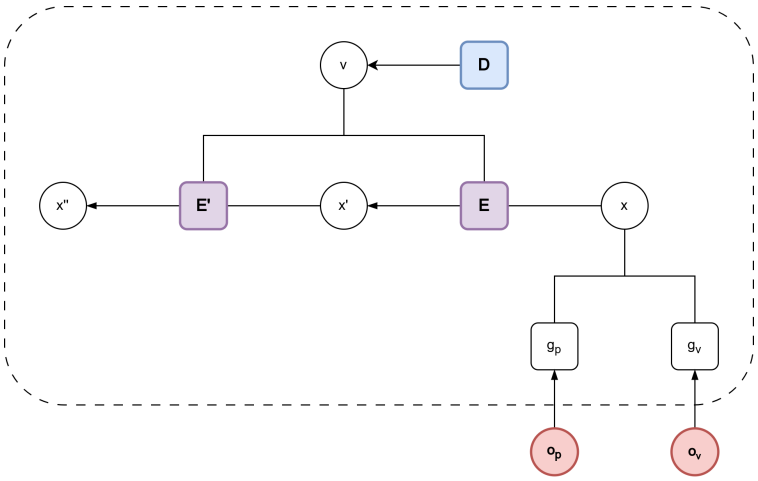


Figure 4: Factor graph for the trajectory inference task. A belief $\boldsymbol{x}$ over the arm and the targets – encoded in joint angles – compute proprioceptive and visual predictions for each component through the functions $\boldsymbol{g}_p$ and $\boldsymbol{g}_v$, respectively. Through the free energies $\boldsymbol{E}$ and $\boldsymbol{E}'$, the most likely agent's intentions $\boldsymbol{v}$ are dynamically inferred.

The prediction errors are then computed by comparing predictions with the corresponding observations:

$$\varepsilon_p = o_p - g_p(\boldsymbol{\mu}) \qquad\qquad \boldsymbol{\varepsilon}_v = \boldsymbol{o}_v - \boldsymbol{g}_v(\boldsymbol{\mu}) \tag{18}$$

On the other hand, the belief over the hidden causes $\boldsymbol{\nu}$ – generated from a categorical distribution with a uniform prior – has only two elements, representing the agent's intentions to follow the first or the second target:

$$\boldsymbol{\nu} = [\nu_{t1} \quad \nu_{t2}] \tag{19}$$

The belief dynamics can be quite complex and take into account many factors including multiple temporal orders – friction, gravity, and so on – but in the following we assume that it depends on a combination of these intentions and that the agent is velocity-controlled. We thus define two functions that set the arm's joint angle equal to the inferred target configurations:

$$\boldsymbol{\mu}^*_{t1} = \boldsymbol{i}_{t1}(\boldsymbol{\mu}) = [\mu_{t1} \quad \mu_{t1} \quad \mu_{t2}] \qquad \boldsymbol{\mu}^*_{t2} = \boldsymbol{i}_{t2}(\boldsymbol{\mu}) = [\mu_{t2} \quad \mu_{t1} \quad \mu_{t2}] \tag{20}$$

Next, we define a reduced dynamics function for each intention, acting as an attractive force proportional to the error between the current and desired beliefs:

$$\begin{aligned} \boldsymbol{f}_{t1}(\boldsymbol{\mu}) &= \lambda \boldsymbol{e}_{t1} & \boldsymbol{e}_{t1} &= \boldsymbol{\mu}^*_{t1} - \boldsymbol{\mu} \\ \boldsymbol{f}_{t2}(\boldsymbol{\mu}) &= \lambda \boldsymbol{e}_{t2} & \boldsymbol{e}_{t2} &= \boldsymbol{\mu}^*_{t2} - \boldsymbol{\mu} \end{aligned} \tag{21}$$

where λ is an attractor gain.

In summary, if the reduced dynamics functions for each object $\boldsymbol{f}_v$ encodes the agent's intention to reach it – which translates to setting the body's component equal to the inferred object configuration – this will generate a trajectory pointing toward that object (i.e., the agents thinks that its arm will be pulled toward the inferred target position). The average velocity is found by weighting the reduced dynamics functions with the corresponding hidden causes – as in Equation 10 – and it is then used to compute a dynamics prediction error:

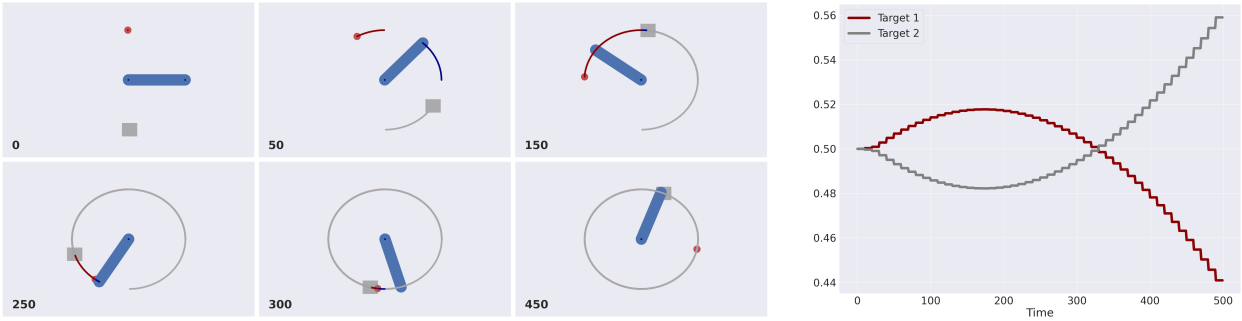
$$\begin{aligned} \boldsymbol{\eta}' &= \nu_{t1} \boldsymbol{f}_{t1}(\boldsymbol{\mu}) + \nu_{t1} \boldsymbol{f}_{t1}(\boldsymbol{\mu}) \\ \boldsymbol{\varepsilon}_x &= \boldsymbol{\mu}' - \boldsymbol{\eta}' \end{aligned} \tag{22}$$

Finally, we can express the update for the hidden states:

$$\dot{\tilde{\boldsymbol{\mu}}} = \mathcal{D}\tilde{\boldsymbol{\mu}} - \partial_{\tilde{\boldsymbol{\mu}}}\mathcal{F} = \begin{bmatrix} \boldsymbol{\mu}' + \partial g_p^T \pi_p \varepsilon_p + \partial \boldsymbol{g}_v^T \pi_v \boldsymbol{\varepsilon}_v + \partial \boldsymbol{\eta}'^T \pi_x \boldsymbol{\varepsilon}_x \\ -\pi_x \boldsymbol{\varepsilon}_x \end{bmatrix} \tag{23}$$

On the other hand, the trajectory inference is done through Equations 11 and 12, where the posteriors over the reduced models are computed by Equation 42. Then, the corresponding reduced log evidence L_v will score how much a possible future trajectory could explain the one that the agent is currently perceiving and realizing. With an uninformative prior, the hidden causes $\boldsymbol{v}$ will then accumulate the evidence for the whole period in which the objects are moving and infer the correct probabilities for each agent's intention.

Finally note that, in the model shown in Figure 4, both orders of the hidden states depend on the same hidden cause: hence, if 2nd-order dynamics and proprioceptive observation are defined, the agent could take into account also the object's acceleration for inferring the same intentions through the free energy $\boldsymbol{E}'$.



(a) (b)

Figure 5: A simple 1-DoF inference task in which the goal is to infer the intention of an agent reaching one of two moving targets. (a) A sample sequence of time frames. The blue arm, the first target (red circle), and the second target (grey square) move around the origin with constant but different velocities. Their trajectories are displayed, too. (b) Probability of hidden causes over time for the two targets. The steps of the trajectories are due to the window of continuous evidence accumulation, which was set to 10 continuous time steps.

Figure 5 illustrates a practical application of this model. The arm and the two targets move at constant angular velocities. However, the red circle and the grey square move slower and faster than the arm, respectively (see Figure 5a). As shown in Figure 5b, both hidden causes are initialized with equal probabilities, but the first cause slowly increases while the arm approaches the first target's position. Then, it reaches a maximum, after which it starts to slowly decrease until the second cause takes over. Overall, this simple task demonstrates that the model can discriminate between causes that generate dynamic trajectories.

2.2.2 Object Grasping

Grasping a moving object (whose graphical model is illustrated in Figure 6) requires more complex computations. In this case, a high-level model performs planning with discrete states $\boldsymbol{s}$ encoding: (i) whether or not the hand is at the object's position; (ii) whether the hand is open or closed; (iii) whether the hand has grasped the object or not. In total, these factors combine in 8 possible process states. These states simultaneously generate discrete predictions for the hidden causes of intrinsic and extrinsic modalities through likelihood matrices $\boldsymbol{A}_i$ and $\boldsymbol{A}_e$. As explained in [31], maintaining a belief over extrinsic information $\boldsymbol{x}_e$ greatly simplifies the computations because it allows to express Cartesian dynamics without worrying about the inverse mapping to the joint angles, which is achieved by inverting through inference a forward kinematic model $\boldsymbol{g}_e$. With this architecture, we can easily express grasping and reaching intentions with intrinsic (corresponding to the hand status *open* or *closed*) and extrinsic (corresponding to the hand position) dynamics, respectively. Nonetheless, more advanced movements – such as grasping an object with a specific wrist rotation – can be achieved by expressing reaching intentions also in the intrinsic modality [32]. Last, a matrix $\boldsymbol{A}_t$ returns a discrete tactile observation $\boldsymbol{o}_t$ – encoded by a Bernoulli distribution – signaling whether or not the target is grasped. In summary:

$$\begin{aligned} \boldsymbol{v}_i &= \boldsymbol{A}_i \boldsymbol{s} = [\boldsymbol{v}_{i,open} \quad \boldsymbol{v}_{i,close}] \\ \boldsymbol{v}_e &= \boldsymbol{A}_e \boldsymbol{s} = [\boldsymbol{v}_{e,stay} \quad \boldsymbol{v}_{e,reach}] \\ \boldsymbol{o}_t &= \boldsymbol{A}_t \boldsymbol{s} = [\boldsymbol{o}_{t,picked} \quad \boldsymbol{o}_{t,placed}] \end{aligned} \tag{24}$$

As regards the discrete model, its actions $\boldsymbol{u}$ correspond to the three steps of the task, plus a *stay* action: (i) open the hand; (ii) reach the target; (iii) close the hand. Note that in this simple implementation, the hidden causes correspond to the discrete actions, hence for simplicity we used the same notation. However, since they are generated from the discrete states, they will generally correspond to more abstract continuous trajectories. Finally, the transition matrix $\boldsymbol{B}$ is defined such that the object can be grasped only when both hand and object are at the same position.

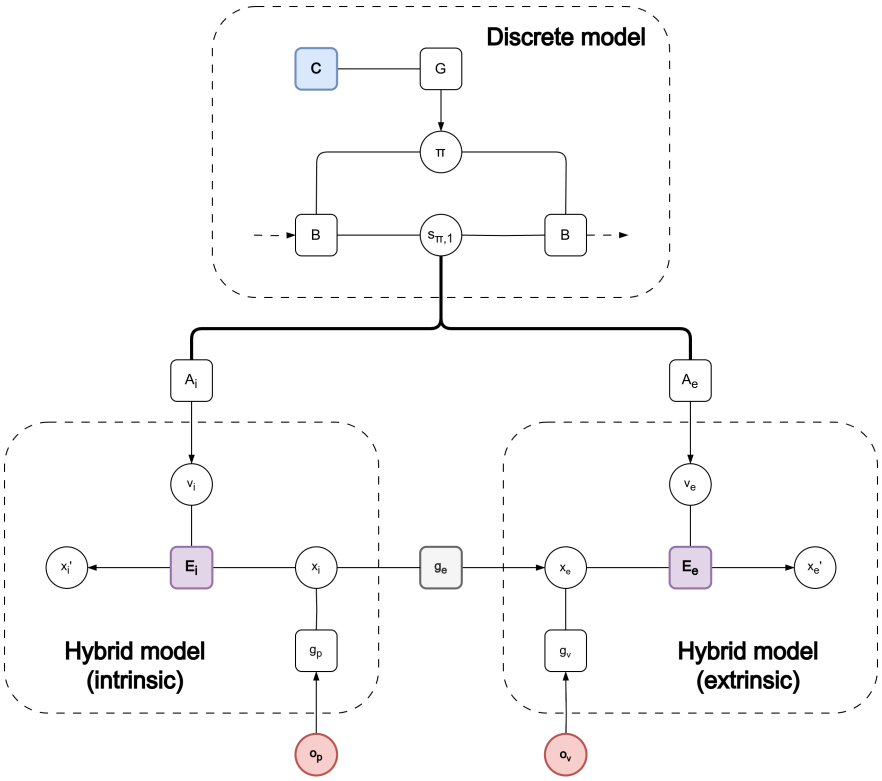


Figure 6: Factor graph for the object grasping task. A discrete state $\boldsymbol{s}_\pi$ at time $\tau = 1$ computes in parallel two predictions of hidden causes $\boldsymbol{v}_i$ and $\boldsymbol{v}_e$ for the lower-level intrinsic and extrinsic modalities through likelihood matrices $\boldsymbol{A}_i$ and $\boldsymbol{A}_e$. In turn, these causes bias the hidden states below $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ generating a particular trajectory. The two modalities are linked by a kinematic generative model $\boldsymbol{g}_e$ and they generate proprioceptive and visual predictions, respectively.

The agent's body is an 8-DoF arm where the last four joints correspond to the fingers, and with the same continuous observations of the previous example:

$$\boldsymbol{o} = \{\boldsymbol{o}_p, \boldsymbol{o}_v\} \tag{25}$$

Visual features are approximated with Cartesian coordinates but may be inferred from a rich visual input [30]. The agent maintains latent states over intrinsic and extrinsic modalities:

$$\boldsymbol{\mu} = \{\boldsymbol{\mu}_i, \boldsymbol{\mu}_e\} \tag{26}$$

that comprise components corresponding to the arm and object, reflecting the decomposition of the visual sensory signal:

$$\boldsymbol{\mu}_i = [\boldsymbol{\mu}_{i,a} \quad \boldsymbol{\mu}_{i,o}] \qquad\qquad \boldsymbol{\mu}_e = [\boldsymbol{\mu}_{e,a} \quad \boldsymbol{\mu}_{e,o}] \tag{27}$$

The kinematic generative model is used for the mapping between intrinsic and extrinsic beliefs:

$$\boldsymbol{p}_e = \boldsymbol{g}_e(\boldsymbol{\mu}_i) = [\boldsymbol{J}(\boldsymbol{\mu}_{i,a}) \quad \boldsymbol{J}(\boldsymbol{\mu}_{i,o}))] + \boldsymbol{w}_e \tag{28}$$

Since we are not interested in the fingers' positions during the reaching task, these functions only compute the Cartesian position of the middle of the two fingers. Now, we define an intention for each hidden cause – how these intentions are defined is explained in [31]. Thus, for the intrinsic modality:

$$\begin{aligned} \boldsymbol{\mu}^*_{i,open} &= \boldsymbol{i}_{i,open}(\boldsymbol{\mu}_i) = \left[\boldsymbol{h}^*_{i,open} \quad \boldsymbol{\mu}_{i,o}\right] \\ \boldsymbol{\mu}^*_{o,close} &= \boldsymbol{i}_{i,close}(\boldsymbol{\mu}_i) = \left[\boldsymbol{h}^*_{i,close} \quad \boldsymbol{\mu}_{i,o}\right] \end{aligned} \tag{29}$$

where:

$$\begin{aligned} \boldsymbol{h}^*_{i,open} &= [\mu_{i,a,1} \quad \dots \quad \mu_{i,a,4} \quad 0 \quad 0 \quad -\theta_c \quad \theta_c] \\ \boldsymbol{h}^*_{i,close} &= [\mu_{i,a,1} \quad \dots \quad \mu_{i,a,4} \quad \theta_o \quad -\theta_o \quad 0 \quad 0] \end{aligned} \tag{30}$$

and θ_c and θ_o are the closed/open angles. Note that, as explained before, the object component of the intrinsic belief $\boldsymbol{\mu}_{i,o}$ is not used here. Correspondingly, for the extrinsic modality:

$$\begin{aligned} \boldsymbol{\mu}^*_{e,stay} &= \boldsymbol{i}_{e,stay}(\boldsymbol{\mu}_e) = \boldsymbol{\mu}_e \\ \boldsymbol{\mu}^*_{e,reach} &= \boldsymbol{i}_{e,reach}(\boldsymbol{\mu}_e) = [\boldsymbol{\mu}_{e,o} \quad \boldsymbol{\mu}_{e,o}] \end{aligned} \tag{31}$$

Note that $\boldsymbol{i}_{e,stay}$ predicts the same belief and can be seen as an intention to maintain the current state of the world.

As before, we now associate a reduced dynamics function – depending on the defined intentions – to a hidden cause of both modalities, which is predicted by the discrete model. The update of the hidden states follows the derivations of the previous example:

$$\begin{aligned} \dot{\tilde{\boldsymbol{\mu}}}_i &= \begin{bmatrix} \boldsymbol{\mu}'_i + \partial \boldsymbol{g}_p^T \pi_p \boldsymbol{\varepsilon}_p + \partial \boldsymbol{\eta}_i'^T \pi_x \boldsymbol{\varepsilon}_{x,i} \\ -\pi_{x,i} \boldsymbol{\varepsilon}_{x,i} \end{bmatrix} \\ \dot{\tilde{\boldsymbol{\mu}}}_e &= \begin{bmatrix} \boldsymbol{\mu}'_e + \partial \boldsymbol{g}_e^T \pi_e \boldsymbol{\varepsilon}_e + \partial \boldsymbol{g}_v^T \pi_v \boldsymbol{\varepsilon}_v + \partial \boldsymbol{\eta}_e'^T \pi_{x,e} \boldsymbol{\varepsilon}_{x,e} \\ -\pi_{x,e} \boldsymbol{\varepsilon}_{x,e} \end{bmatrix} \end{aligned} \tag{32}$$

As concerns the hidden causes, they are updated by averaging the high-level expectation and the low-level evidence:

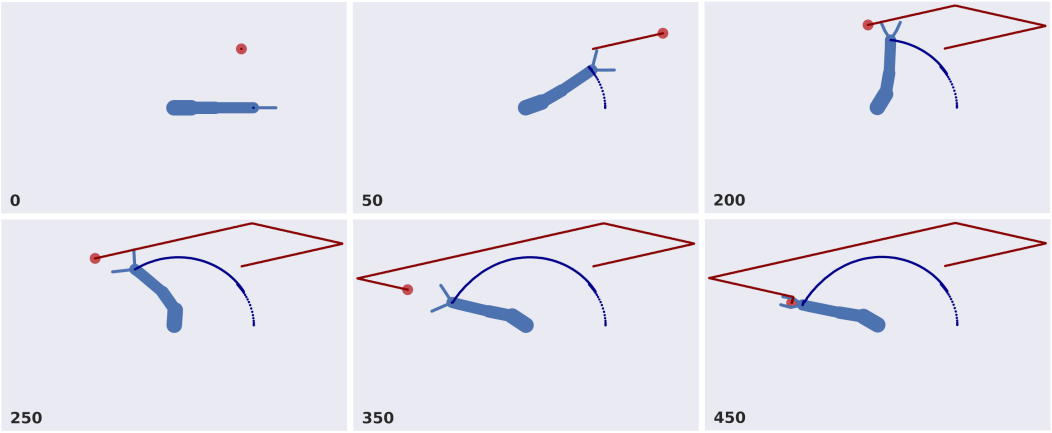
$$\begin{aligned} \boldsymbol{\nu}_i &= \sigma(-\boldsymbol{E}_i) = \sigma(\ln \boldsymbol{A}_i \boldsymbol{s} + \int_0^T \boldsymbol{\mathcal{L}}_i dt) \\ \boldsymbol{\nu}_e &= \sigma(-\boldsymbol{E}_e) = \sigma(\ln \boldsymbol{A}_e \boldsymbol{s} + \int_0^T \boldsymbol{\mathcal{L}}_e dt) \end{aligned} \tag{33}$$

These messages are finally used by the discrete model, along with the discrete observation, to infer the most probable discrete hidden state at time τ and conditioned over the policy $\boldsymbol{\pi}$:

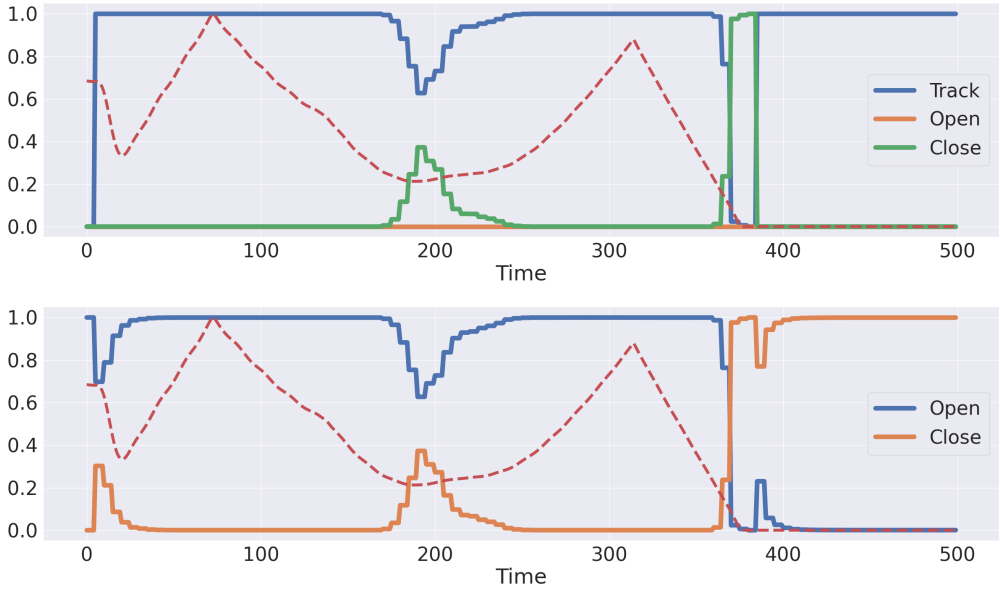
$$\boldsymbol{s}_{\pi,\tau} = \sigma(\ln \boldsymbol{B}_{\pi,\tau-1} \boldsymbol{s}_{\tau-1} + \ln \boldsymbol{A}_i^T \boldsymbol{\nu}_{i,\tau} + \ln \boldsymbol{A}_e^T \boldsymbol{\nu}_{e,\tau} + \ln \boldsymbol{A}_t^T \boldsymbol{o}_t) \tag{34}$$

Note that the evidence accumulation is performed within the low levels, so that the higher-level model has only to compute discrete predictions for each modality and retrieve the evidence by weighting the information from both lower-level models.

Differently from the first example, the reduced free energies $E_{i,v}$ and $E_{e,v}$ will depend both on low-level evidences that signal whether a reaching or grasping intention has been realized, and on high-level expectations that impose the next



(a)



(b)

Figure 7: Object grasping with a 4-DoF arm. (A) A sample sequence of time frames showing the moving target (red dot) and the blue arm performing the grasping action. Trajectories with their respective colors are shown, too. (B) Discrete hidden states (top panel) and continuous intrinsic hidden causes (bottom panel) over time (see legends). Both panels also show the distance between the hand and the object (dashed red line), normalized between zero and one.

discrete step and ultimately arise from the computation of the expected free energy at the top. As a result, if a grasping action fails because the object is moving too fast, the bottom-up message from the hidden causes to the discrete hidden states would inform the discrete model that the hand turned away from the object, and would let it recompute a reaching movement.

This behavior can be seen in Figure 7b displaying the trajectory of discrete hidden states $\boldsymbol{s}$ and intrinsic hidden causes $\boldsymbol{v}_i$. As soon as the hand approaches the object (at t=200), the model makes a first attempt at grasping, resulting in a gradual decrease of the tracking state and a gradual increase of the closed hand state. These states generate hidden causes in both modalities so that the *open* and *close* continuous intentions will change accordingly. However, since the object is still too far from the hand, the discrete states rapidly turn over to tracking. At about t=400, the object is close enough to the hand, resulting in a much steeper change of discrete states and continuous causes. Eventually, the object is correctly grasped.

3 Discussion

Bayesian Model Reduction is a powerful technique that, among its wide usages, allows shrinking the infinite amount of values that a continuous signal can assume into a small set of hypotheses, which in turn can be associated with a few discrete categories for high-level decision-making. How these hypotheses are defined is however crucial if one wants to model an agent with planning capabilities in a highly dynamic environment such as the real world.

Using a-priori defined assumptions can be appropriate for a static environment, e.g., in case the agent is required to reach multiple fixed positions [1]. In this case, expected observations generated from a discrete agent are combined with reduced models to compute a continuous prior for the hidden causes below. Bayesian Model Reduction is then used to compare the posterior over the hidden causes with the defined reduced models, in order to find the most likely observation for the discrete agent. This approach can be used, for example, to simulate pictographic reading [24], where the center of fixation of the agent is determined by a sentence explaining the current context, and through which evidence about a pictogram is accumulated.

Despite the capabilities of this approach, dealing with dynamic contexts – e.g., if one has to reach or grasp moving objects – is much more challenging. A proposal on how to efficiently implement goal-directed behavior in a dynamic context was presented in [30], where the dynamics function of the agent was decomposed into independent *intentions*, each one pulling the belief toward a particular state. Concurrently maintaining such intentions – e.g., reaching different objects – allowed two advanced behaviors: (i) performing a goal-directed movement that results from a combination of different intentions [31]; (ii) implementing a sort of fixed multi-step behavior, e.g., tracking two objects where the toggle event is triggered by a tactile sensation [33], or performing a dynamic pick-and-place operation [32]. These flexible behaviors were implemented through the same approach, i.e., by dynamically modulating the precisions of the dynamics functions where the intentions were embedded. Although having interesting implications – as we also mention here – and being a convenient approach for artificial agents, from a biological perspective it is unclear to what extent the learning of such precisions could be fast enough to account for complex behavior in a highly dynamic context.

A more powerful solution to dynamic multi-step tasks that also affords efficient planning was presented in [32]. This alternative approach makes use of a conventional hybrid model and it is similar to state-of-the-art models such as the one mentioned above. Dynamic behavior in this case was achieved by continuously updating, at each discrete step, the reduced priors of the discrete model depending on the continuous hidden states. However, in this simplified model the update was performed through a direct function, while a more realistic setting would involve the inference and learning of the reduced priors from the generated prediction errors.

Here, we propose that, instead of using reduced models with static priors and performing a BMR with the posterior over the hidden causes, a more efficient, fluent, and biologically plausible behavior can be achieved if we consider the hidden causes as encoding a categorical distribution and if the BMR is performed between the temporal orders of the hidden states. As a result, the reduced priors would depend on dynamics functions that specify how the hidden states could evolve in the future, while the latter are continuously updated through sensory evidence. If we consider the hidden causes as encoding the agent's intentions, an interesting parallelism arises with the policies of a discrete model. The top-down BMA would represent the average intention that the agent wants to realize, while the bottom-up BMC would convey the most likely intention that could have generated the current trajectory, that is, it would signal whether an intention has been realized or a future state has been reached.

With our novel approach, dynamic evidence can be easily accumulated and associated with fixed intentions, which can be useful in every kind of context that presents dynamic elements, e.g., if one needs to infer the trajectory of another agent, or solve a multi-step task such as picking, moving, and releasing an object. In the latter case, the dynamic evidence accumulation is performed within the lower levels, so that the high level needs only to receive a discrete bottom-up message, infer the most probable state, perform planning operations, and finally send intention predictions back to the low levels, thus synchronizing their behavior.

We show that two kinds of inferential processes emerge concerning the reduced dynamics functions. A fast process involving both the inference of the current trajectory through the gradient of the full prior with respect to the hidden states, and the inference of the agent's intentions encoded in the hidden causes. But a slow process exists as well: the inference – or better, learning – of the precisions of the dynamics functions. We propose that such precisions may have a similar interpretation to the confidence assignment mechanisms of the sensory domains: if an intention is not useful for minimizing the prediction errors for a given context, its precision will gradually decrease over time because the agent is not confident about using it anymore – either for realizing its goals or for understanding the state of affairs of the world.

Overall, the model that we presented might explain where and how the reduced priors are encoded in a biologically plausible way. But it might also explain how intentions are assigned to specific tasks and how they emerge from high-level computations. As shown in Figure 7b, although we used elementary discrete actions of reaching and grasping,

a smooth approaching transition occurs that is due to the BMR between the dynamic reduced priors. This novel behavior could be mapped into the continuous dynamics and called in a single action by the discrete model, producing a highly flexible and fluid behavior, as simulated in [32].

Future analyses are needed to simulate these task-specialization and transition mechanisms. One limitation of this study is that we kept the dynamics precisions fixed throughout all the simulations. Letting them adapt and depend on the free energy – as for the hidden states and causes – would be a promising direction of research to simulate how an agent can learn different continuous intentions from a discrete model. To this concern, an interesting approach was implemented in [34], which used multiple continuous generative models followed by a "switcher". Although not using a standard hybrid model, this approach has some analogies with the model presented here in that the agent – in that case, a student bird – can maintain models of how teachers generate a possible evolution of sensory signals, and through the switcher it can perform online model selection to infer the most likely cause that could have generated what it is perceiving.

4 Methods

4.1 Bayesian Model Reduction

Consider a generative model $p(\boldsymbol{\theta}, \boldsymbol{y})$ with parameters $\boldsymbol{\theta}$ and data $\boldsymbol{y}$:

$$p(\boldsymbol{\theta}, \boldsymbol{y}) = p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{35}$$

Consider then an additional distribution $p(\boldsymbol{y}, \boldsymbol{\theta}|m)$ which is a reduced version of the first model if the likelihood of some data is the same under both models and the only difference rests upon the specification of the priors $p(\boldsymbol{\theta}|m)$. We can express the posterior of the reduced model in terms of the posterior of the full model and the ratios of the priors and the evidence:

$$p(\boldsymbol{\theta}|\boldsymbol{y}, m) = p(\boldsymbol{\theta}|\boldsymbol{y})\frac{p(\boldsymbol{\theta}|m)p(\boldsymbol{y})}{p(\boldsymbol{\theta})p(\boldsymbol{y}|m)} \tag{36}$$

The procedure for computing the reduced posterior is the following: first, we integrate over the parameters to obtain the evidence ratio of the two models:

$$p(\boldsymbol{y}|m) = p(\boldsymbol{y}) \int p(\boldsymbol{\theta}|\boldsymbol{y})\frac{p(\boldsymbol{\theta}|m)}{p(\boldsymbol{\theta})}d\boldsymbol{\theta} \tag{37}$$

Then, we define an approximate posterior $q(\boldsymbol{\theta})$ and we compute the reduced Variational Free Energy (VFE) in terms of the full model:

$$\mathcal{F}[p(\boldsymbol{\theta}|m)] \approx \mathcal{F}[p(\boldsymbol{\theta})] + \ln \mathop{\mathbb{E}}_{q}\left[\frac{p(\boldsymbol{\theta}|m)}{p(\boldsymbol{\theta})}\right] \tag{38}$$

In other words, this quantity acts as a hint to how well the reduced representation explains the full model Similarly, the approximate posterior of the reduced model can be written in terms of the posterior of the full model:

$$\ln q(\boldsymbol{\theta}|m) = \ln q(\boldsymbol{\theta}) + \ln \frac{p(\boldsymbol{\theta}|m)}{p(\boldsymbol{\theta})} - \ln \mathop{\mathbb{E}}_{q}\left[\frac{p(\boldsymbol{\theta}|m)}{p(\boldsymbol{\theta})}\right] \tag{39}$$

The Laplace approximation [35] leads to a simple form of the approximate posterior and the reduced free energy. Assuming the following Gaussian distributions:

$$\begin{aligned} p(\boldsymbol{\theta}) &= \mathcal{N}(\boldsymbol{\eta}, \boldsymbol{\Sigma}) & q(\boldsymbol{\theta}) &= \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C}) \\ p(\boldsymbol{\theta}|m) &= \mathcal{N}(\boldsymbol{\eta}_m, \boldsymbol{\Sigma}_m) & q(\boldsymbol{\theta}|m) &= \mathcal{N}(\boldsymbol{\mu}_m, \boldsymbol{C}_m) \end{aligned} \tag{40}$$

The reduced free energy turns to:

$$\mathcal{F}[p(\boldsymbol{\theta}|m)] \approx \frac{1}{2}\ln|\boldsymbol{\Pi}_m \boldsymbol{P} \boldsymbol{C}_m \boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{\mu}^T \boldsymbol{P} \boldsymbol{\mu} - \boldsymbol{\mu}_m^T \boldsymbol{P}_m \boldsymbol{\mu}_m - \boldsymbol{\eta}^T \boldsymbol{\Pi} \boldsymbol{\eta} + \boldsymbol{\eta}_m^T \boldsymbol{\Pi}_m \boldsymbol{\eta}_m) \tag{41}$$

expressed in terms of precision of priors $\boldsymbol{P}$ and posteriors $\boldsymbol{\Pi}$. The reduced posterior mean and precision are computed through Equation 39:

$$\begin{aligned} \boldsymbol{\mu}_m &= \boldsymbol{C}_m(\boldsymbol{P}\boldsymbol{\mu} - \boldsymbol{\Pi}\boldsymbol{\eta} + \boldsymbol{\Pi}_m \boldsymbol{\eta}_m) \\ \boldsymbol{P}_m &= \boldsymbol{P} - \boldsymbol{\Pi} + \boldsymbol{\Pi}_m \end{aligned} \tag{42}$$

4.2 Active Inference in Continuous Time

Active Inference operates under the assumption that organisms perceive their environment by employing an internal generative model, inferring how external causes lead to sensory signals within the real generative process [4, 2]. Within the continuous domain, this procedure is commonly broken down into probability distributions encompassing hidden causes $\tilde{\boldsymbol{v}}$, hidden states $\tilde{\boldsymbol{x}}$, and observable outcomes $\tilde{\boldsymbol{y}}$:

$$p(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}, \tilde{\boldsymbol{y}}) = p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}})p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}})p(\tilde{\boldsymbol{v}}) \tag{43}$$

These distributions are approximated through Gaussian functions:

$$\begin{aligned} p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{\Pi}}_y^{-1}) \\ p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}}) &= \mathcal{N}(\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}), \tilde{\boldsymbol{\Pi}}_x^{-1}) \\ p(\tilde{\boldsymbol{v}}|\boldsymbol{\eta}) &= \mathcal{N}(\boldsymbol{\eta}, \tilde{\boldsymbol{\Pi}}_v^{-1}) \end{aligned} \tag{44}$$

These correspond to the subsequent nonlinear stochastic equations that describe the evolution of the environment:

$$\begin{aligned} \tilde{\boldsymbol{y}} &= \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}) + \boldsymbol{w}_y \\ \mathcal{D}\tilde{\boldsymbol{x}} &= \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}) + \boldsymbol{w}_x \end{aligned} \tag{45}$$

Here, $\mathcal{D}$ is a differential shift operator, while the symbol $\sim$ denotes variables encoded in generalized coordinates that represent their instantaneous trajectories.

The direct computation of the posterior $p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})$ is unfeasible, as it involves the calculation of the inaccessible marginal $p(\tilde{\boldsymbol{y}})$. The variational approach involves approximating the posteriors with manageable distributions, such as Gaussians:

$$\begin{aligned} q(\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{x}}; \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{P}}_x^{-1}) \\ q(\tilde{\boldsymbol{v}}) &= \mathcal{N}(\tilde{\boldsymbol{v}}; \tilde{\boldsymbol{\nu}}, \tilde{\boldsymbol{P}}_v^{-1}) \end{aligned} \tag{46}$$

Where the parameters $\tilde{\boldsymbol{\mu}}$ and $\tilde{\boldsymbol{\nu}}$ are termed beliefs concerning hidden states and hidden causes, respectively. The process of perceptual inference becomes a minimization, aiming to reduce the disparity between the approximated and actual posteriors. This can be framed in terms of the KL divergence – equivalent to the expectation, over the approximated posterior, of the discrepancy between the two logarithmic probabilities:

$$DKL[q(\tilde{\boldsymbol{x}})||p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})] = \sum\nolimits_x q(\tilde{\boldsymbol{x}}) \ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})} = \mathbb{E}\, q(\tilde{\boldsymbol{x}})[\ln q(\tilde{\boldsymbol{x}}) - \ln p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})] \tag{47}$$

As the denominator $p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})$ still hinges on the marginal $p(\tilde{\boldsymbol{y}})$, the KL divergence is expressed in terms of the log-evidence and the *Variational Free Energy* (VFE), with the latter being minimized. Given that the KL divergence is non-negative (due to Jensen's inequality), the VFE serves as an upper bound on the log-evidence. Consequently, minimizing the VFE enhances both the evidence and model fit:

$$\mathcal{F} = \mathbb{E}\, q(\tilde{\boldsymbol{x}}) \left[\ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})}\right] = \underset{q(\tilde{\boldsymbol{x}})}{\mathbb{E}} \left[\ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})}\right] - \ln p(\tilde{\boldsymbol{y}}) \tag{48}$$

Within the given approximations, the minimization of the free energy involves an iterative update of parameters through the following expressions:

$$\begin{aligned} \dot{\tilde{\boldsymbol{\mu}}} - \mathcal{D}\tilde{\boldsymbol{\mu}} &= -\partial_{\boldsymbol{\mu}}\mathcal{F} = \partial\tilde{\boldsymbol{g}}^T\tilde{\boldsymbol{\Pi}}_y\tilde{\boldsymbol{\varepsilon}}_y + \partial_{\boldsymbol{\mu}}\tilde{\boldsymbol{f}}^T\tilde{\boldsymbol{\Pi}}_x\tilde{\boldsymbol{\varepsilon}}_x - \mathcal{D}^T\tilde{\boldsymbol{\Pi}}_x\tilde{\boldsymbol{\varepsilon}}_x \\ \dot{\tilde{\boldsymbol{\nu}}} - \mathcal{D}\tilde{\boldsymbol{\nu}} &= -\partial_{\boldsymbol{\nu}}\mathcal{F} = \partial_{\boldsymbol{\nu}}\tilde{\boldsymbol{f}}^T\tilde{\boldsymbol{\Pi}}_x\tilde{\boldsymbol{\varepsilon}}_x - \tilde{\boldsymbol{\Pi}}_v\tilde{\boldsymbol{\varepsilon}}_v \end{aligned} \tag{49}$$

Here, $\tilde{\boldsymbol{\varepsilon}}_s$, $\tilde{\boldsymbol{\varepsilon}}_x$, and $\tilde{\boldsymbol{\varepsilon}}_v$ respectively stand for prediction errors in sensations, dynamics, and prior information:

$$\begin{aligned} \tilde{\boldsymbol{\varepsilon}}_y &= \tilde{\boldsymbol{y}} - \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{\mu}}) \\ \tilde{\boldsymbol{\varepsilon}}_x &= \mathcal{D}\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\nu}}) \\ \tilde{\boldsymbol{\varepsilon}}_v &= \tilde{\boldsymbol{\nu}} - \boldsymbol{\eta} \end{aligned} \tag{50}$$

Importantly, the VFE can also be employed to compute motor control signals:

$$\dot{\boldsymbol{a}} = -\partial_{\boldsymbol{a}}\mathcal{F} = -\partial_{\boldsymbol{a}}\tilde{\boldsymbol{g}}^T\tilde{\boldsymbol{\Pi}}_y\tilde{\boldsymbol{\varepsilon}}_y \tag{51}$$

This process facilitates the selection of sensations corresponding to the current representation of the world by the agent, which embodies Active Inference. This sequence further diminishes the VFE, enabling the implementation of goal-directed behavior and maintaining the agent within predictable and safer spaces [4].

4.3 Active Inference in Discrete Time

While Active Inference in continuous time can address real-world challenges by tracking the instantaneous trajectories generated by the continuous environment, it does come with several limitations, constraining its applicability as it struggles to accommodate broader classes of actions, including decision-making. The VFE minimization can solely adapt the approximate posterior based on current or past observations, failing to incorporate possible future states and outcomes. To confer this capability upon an agent, a concept called the *Expected Free Energy* (EFE) is introduced [12, 13]. The initial step involves augmenting the generative model with policies $\boldsymbol{\pi}$:

$$p(\boldsymbol{s}, \boldsymbol{o}, \boldsymbol{\pi}) = p(\boldsymbol{o}|\boldsymbol{s}, \boldsymbol{\pi})p(\boldsymbol{s}|\boldsymbol{\pi})p(\boldsymbol{\pi}) \tag{52}$$

Here, $\boldsymbol{s}$ and $\boldsymbol{o}$ denote discrete states and outcomes. Notably, these policies are not mere stimulus-response mappings like those in Reinforcement Learning schemes; rather, they encompass sequences of actions. Given that action planning necessitates the selection of policies leading to desired priors, future outcomes not yet observed must also be taken into account. Hence, the EFE is formulated by conditioning over these unobserved outcomes, treating them as hidden states:

$$\mathcal{G}\pi = \mathbb{E}\, q(\boldsymbol{s}, \boldsymbol{o}|\boldsymbol{\pi}) \left[\ln \frac{q(\boldsymbol{s}|\boldsymbol{\pi})}{p(\boldsymbol{s}, \boldsymbol{o}|\boldsymbol{\pi})}\right] \approx \underset{q(\boldsymbol{s},\boldsymbol{o}|\boldsymbol{\pi})}{\mathbb{E}} \left[\ln \frac{q(\boldsymbol{s})}{q(\boldsymbol{s}|\boldsymbol{o}, \boldsymbol{\pi})}\right] - \underset{q(\boldsymbol{o}|\boldsymbol{\pi})}{\mathbb{E}} [\ln p(\boldsymbol{o}|\boldsymbol{C})] \tag{53}$$

Here, the probability distribution $p(\boldsymbol{o}|\boldsymbol{C})$ represents preferred outcomes. The last two terms are referred to as *epistemic* (uncertainty-reducing) and *pragmatic* (goal-seeking) terms, respectively. In practical terms, this quantity is employed by initially factoring the agent's generative model similar to POMDPs:

$$p(\boldsymbol{s}_{1:T}, \boldsymbol{o}_{1:T}, \boldsymbol{\pi}) = p(\boldsymbol{s}_1) \cdot p(\boldsymbol{\pi}) \cdot \prod_{\tau=1}^{T} p(\boldsymbol{o}_\tau|\boldsymbol{s}_\tau) \cdot \prod_{\tau=2}^{T} p(\boldsymbol{s}_\tau|\boldsymbol{s}_{\tau-1}, \boldsymbol{\pi}) \tag{54}$$

Each of these elements can be expressed using categorical distributions:

$$\begin{aligned} p(\boldsymbol{s}_1) &= Cat(\boldsymbol{D}) & p(\boldsymbol{o}_\tau|\boldsymbol{s}_\tau) &= Cat(\boldsymbol{A}) \\ p(\boldsymbol{\pi}) &= Cat(\boldsymbol{E}) & p(\boldsymbol{s}_\tau|\boldsymbol{s}_{\tau-1}, \boldsymbol{\pi}) &= Cat(\boldsymbol{B}_{\pi,\tau}) \end{aligned} \tag{55}$$

Here, $\boldsymbol{D}$ captures beliefs about the initial state, $\boldsymbol{E}$ encodes the prior over policies, $\boldsymbol{A}$ represents the likelihood matrix, and $\boldsymbol{B}_{\pi,\tau}$ is the transition matrix. Assuming a mean-field approximation, where the approximate posterior factorizes into independent distributions, the approach continues:

$$\begin{aligned} p(\boldsymbol{s}_{1:T}|\boldsymbol{o}_{1:T}, \boldsymbol{\pi}) \approx q(\boldsymbol{s}_{1:T}, \boldsymbol{\pi}) &= q(\boldsymbol{\pi}) \prod_{\tau}^{T} q(\boldsymbol{s}_\tau|\pi) \\ q(\boldsymbol{\pi}) &= Cat(\boldsymbol{\pi}) \\ q(\boldsymbol{s}_\tau|\boldsymbol{\pi}) &= Cat(\boldsymbol{s}_{\pi,\tau}) \end{aligned} \tag{56}$$

Variational message passing can then be employed to infer each posterior $q(\boldsymbol{s}_\tau|\boldsymbol{\pi})$ and combine them into a global posterior $q(\boldsymbol{s}_{1:T}|\boldsymbol{\pi})$. To update the posterior about hidden states, the method combines messages from past and future states, as well as outcomes. This entails expressing each term in relation to its sufficient statistics and subsequently applying a softmax function to achieve a valid probability distribution:

$$\boldsymbol{s}_{\pi,\tau} = \sigma(\ln \boldsymbol{B}_{\pi,\tau-1}\boldsymbol{s}_{\pi,\tau-1} + \ln \boldsymbol{B}_{\pi,\tau}^T \boldsymbol{s}_{\pi,\tau+1} + \ln \boldsymbol{A}^T \boldsymbol{o}_\tau) \tag{57}$$

Likewise, to update $q(\boldsymbol{\pi})$, the approach involves combining messages from the policy's prior (expressed by matrix $\boldsymbol{E}$) and from future observations conditioned on policies. The latter can be approximated by the EFE conditioned at a specific time τ:

$$\begin{aligned} \mathcal{G}_\pi &\approx \sum_\tau D_{KL}[q(\boldsymbol{o}_\tau|\boldsymbol{\pi})||p(\boldsymbol{o}_\tau|\boldsymbol{C})] + \underset{q(\boldsymbol{s}_\tau|\boldsymbol{s}_{\tau-1},\boldsymbol{\pi})}{\mathbb{E}} [H[p(\boldsymbol{o}_\tau|\boldsymbol{s}_\tau)]] \\ &= \sum_\tau \boldsymbol{o}_{\pi,\tau}(\ln \boldsymbol{o}_{\pi,\tau} - \boldsymbol{C}_\tau) + \boldsymbol{s}_{\pi,\tau}\boldsymbol{H} \end{aligned} \tag{58}$$

Here:

$$\boldsymbol{o}_{\pi,\tau} = \boldsymbol{A}\boldsymbol{s}_{\pi,\tau} \qquad \boldsymbol{C}_\tau = \ln p(\boldsymbol{o}_\tau|\boldsymbol{C}) \qquad \boldsymbol{H} = -diag(\boldsymbol{A}^T \ln \boldsymbol{A}) \tag{59}$$

Ultimately, the most probable action under all policies is selected:

$$\begin{aligned} \boldsymbol{\pi} &= \sigma(\ln \boldsymbol{E} - \boldsymbol{\mathcal{G}}) \\ u_t &= \arg\max_u \boldsymbol{\pi} \cdot [U_{\pi,t} = u] \end{aligned} \tag{60}$$

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.S. and the Italian Ministry for Research MIUR under Grant Agreement PRIN 2017KZNZLN to I.S. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] Thomas Parr, Giovanni Pezzulo, and Karl J Friston. *Active inference: the free energy principle in mind, brain, and behavior*. Cambridge, MA: MIT Press, 2021.

[2] Karl Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

[3] 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.

[4] 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.

[5] 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.

[6] Harriet Brown, Karl Friston, and Sven Bestmann. Active inference, attention, and motor preparation. *Frontiers in Psychology*, 2(SEP):1–10, 2011.

[7] Karl J. Friston, Klaas Enno Stephan, Baojuan Li, and Jean Daunizeau. Generalised filtering. 2010.

[8] Karl Friston. Hierarchical models in the brain. *PLoS Computational Biology*, 4(11), 2008.

[9] Matthew Botvinick and Marc Toussaint. Planning as inference. *Trends in Cognitive Sciences*, 16(10):485–488, 2012.

[10] Marc Toussaint. Probabilistic inference as a model of planned behavior. *Künstliche Intelligenz*, 3/09:23–29, 2009.

[11] 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.

[12] Lancelot Da Costa, Thomas Parr, Noor Sajid, Sebastijan Veselic, Victorita Neacsu, and Karl Friston. Active inference on discrete state-spaces: A synthesis. *Journal of Mathematical Psychology*, 99, 2020.

[13] 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.

[14] Raphael Kaplan and Karl J. Friston. Planning and navigation as active inference. *Biological Cybernetics*, 112(4):323–343, 2018.

[15] Karl Friston, Thomas Parr, and Peter Zeidman. Bayesian model reduction. pages 1–32, 2018.

[16] Karl Friston and Will Penny. Post hoc Bayesian model selection. *NeuroImage*, 56(4):2089–2099, 2011.

[17] K. J. Friston, L. Harrison, and Will Penny. Dynamic causal modelling. *NeuroImage*, 19(4):1273–1302, 2003.

[18] Anne G. E. Collins and Michael J. Frank. Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychological Review*, 120(1):190–229, 2013.

[19] Karl Friston, Peter Zeidman, and Vladimir Litvak. Empirical bayes for DCM: A group inversion scheme. *Frontiers in Systems Neuroscience*, 9(November):1–10, 2015.

[20] Ruslan Salakhutdinov, Joshua B. Tenenbaum, and Antonio Torralba. Learning with hierarchical-deep models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1958–1971, 2013.

[21] Joshua B. Tenenbaum, Charles Kemp, Thomas L. Griffiths, and Noah D. Goodman. How to grow a mind: Statistics, structure, and abstraction. *Science*, 331(6022):1279–1285, 2011.

[22] D. Gowanlock R. Tervo, Joshua B. Tenenbaum, and Samuel J. Gershman. Toward the neural implementation of structure learning. *Current Opinion in Neurobiology*, 37:99–105, 2016.

[23] Mark W. Woolrich, Saad Jbabdi, Brian Patenaude, Michael Chappell, Salima Makni, Timothy Behrens, Christian Beckmann, Mark Jenkinson, and Stephen M. Smith. Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, 45(1 Suppl):S173–S186, 2009.

[24] Karl J. Friston, Thomas Parr, and Bert de Vries. The graphical brain: Belief propagation and active inference. 1(4):381–414, 2017.

[25] 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.

[26] Thomas Parr and Karl J. Fristo. The Discrete and Continuous Brain: From Decisions to Movement—And Back Again. (September):2319–2347, 2018.

[27] T. Parr and K. J. Friston. The computational pharmacology of oculomotion. *Psychopharmacology (Berl.)*, 236(8):2473–2484, August 2019.

[28] K. J. Friston, N. Sajid, D. R. Quiroga-Martinez, T. Parr, C. J. Price, and E. Holmes. Active listening. *Hear. Res.*, 399:107998, January 2021.

[29] 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.

[30] Matteo Priorelli and Ivilin Peev Stoianov. Flexible intentions: An active inference theory. *Frontiers in Computational Neuroscience*, 17:1128694:1–27, Mar. 2023.

[31] Matteo Priorelli, Giovanni Pezzulo, and Ivilin Peev Stoianov. Deep kinematic inference affords efficient and scalable control of bodily movements. *bioRxiv*, pages 1–33, 2023.

[32] Matteo Priorelli and Ivilin Peev Stoianov. Slow but flexible or fast but rigid? discrete and continuous processes compared. *bioRxiv*, 2023.

[33] Matteo Priorelli and Ivilin Peev Stoianov. Intention Modulation for Multi-Step Tasks in Continuous Time Active Inference. In *Active Inference, Third International Workshop, IWAI 2022, Grenoble, France, Sept 19, 2022*, 2022.

[34] 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.

[35] Karl Friston, Jérémie Mattout, Nelson Trujillo-Barreto, John Ashburner, and Will Penny. Variational free energy and the Laplace approximation. *NeuroImage*, 34(1):220–234, 2007.