DYNAMIC INFERENCE BY MODEL REDUCTION

Matteo Priorelli
Sapienza University of Rome, Italy
Institute of Cognitive Sciences and Technologies,
National Research Council of Italy

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

ABSTRACT

How can agents infer the intentions of others by simply observing their behavior? And how can they generate fast and accurate actions such as grasping a moving object on the fly? Recent advances in Bayesian model reduction have led to innovative, biologically plausible approaches to actively infer the state of affairs of the world and perform planning with continuous signals. However, reducing the surrounding environment into a small set of simpler hypotheses remains a challenge in highly dynamic contexts. In this study, we propose an approach, based on active inference, that employs dynamic priors sampled from reduced versions of a generative model. Each dynamic prior corresponds to an alternative evolution of the world, which the agent can evaluate by accumulating continuous data. We test our approach on two everyday tasks: inferring a trajectory and grasping a moving object. Our findings reveal how agents can smoothly infer and enact dynamic intentions, and emphasize the key role of intentional gain or precision in motor learning.

1 Introduction

How can intelligent agents construct dynamic representations of the world that support high-level planning, control, and decision making? A recent brain-inspired framework, known as *active inference*, proposes that in order to survive, all organisms are constantly involved in minimizing a quantity called *free energy* [1, 2, 3, 4]. To remain within viable states, agents are supposed to maintain an internal generative model about how hidden states generate sensory observations and how such states evolve. The resulting predictions are compared with the observations to form prediction errors, which are minimized either by updating the agent's beliefs (giving rise to perception) or by changing the world through action [5, 6].

Dealing with a continuous environment is possible by expressing the hidden states in terms of generalized coordinates consisting of multiple temporal orders (e.g., position, velocity, acceleration, and so on); this permits representing the environment dynamics with high fidelity [7, 8, 9], but only minimizes the free energy of present states and observations. In other words, simply responding reflexively to changes in sensory input precludes planning. This is sometimes referred to as "merely reflexive" active inference (e.g., as a thermostat or small insect). Decision-making is possible by augmenting a continuous model with a discretized representation of the environment, encoding possible future states. Planning then arises when future outcomes are treated as hidden states – an approach called *planning as inference* [10, 11, 12] – in which actions are selected to minimize the free energy expected in the future in consequence of that action [13, 14, 15]. However, selecting among a number of discrete courses of action (i.e., sequential policies) raises a problem: how does one infer which discrete trajectory is being pursued, given observations of continuous movement. In active inference, this is usually implemented using Bayesian model reduction to evaluate the evidence for a small number of discrete alternatives. Bayesian model reduction [16] is an efficient form of Bayesian model selection that has been used in several areas over the course of the years [17, 18, 19, 20, 21, 22, 23, 24]. This technique consists of reducing a complex distribution into a small set of simpler hypotheses or reduced models that have the same likelihood and differ only in their priors.

Hybrid models that integrate discrete and continuous components have been applied to several domains, ranging from epistemic foraging in reading, [25, 26], multi-step reaching movements in neurological disorders [27], saccades and visual sampling [28, 29], active listening [30], and interoceptive control [31]. However, the reduced models used for evaluating the full complex distribution are usually shaped a priori. This limits the agent to a static hypothesis space, making it difficult to adapt quickly in dynamic environments.

Here, we propose an alternative method that combines discrete hypotheses with potential trajectories. In summary, we address a fundamental problem in neuroscience and neurorobotics; namely, inferring the discrete intentions of another – or indeed oneself – based on patterns of continuous movement. This is relevant, not just in terms of action observation, but also in refining predictions of one's own actions in the context of active inference. Technically, we solve this problem by combining Bayesian model reduction with a mixture of Gaussians over movement trajectories in generalized coordinates of motion. This allows the agent to flexibly construct and evaluate potential futures in real-time, enabling both efficient inference and rapid adaptation in constantly changing settings. This approach can be used in several dynamic contexts, and we demonstrate its capabilities in inferring which – of a number of moving targets – an agent is tracking, and grasping a moving object.

2 Results

2.1 Dynamic hybrid models

As described in Section 4.4, conventional hybrid models in active inference operate under the assumption that discrete hypotheses correspond to fixed priors over continuous latent states. In simple environments with little uncertainty this is often sufficient; however, it becomes limiting in real-world scenarios where intentions unfold dynamically over time, such as inferring a followed target or grasping a moving object.

The first problem can be addressed by considering the target positions as priors of reduced models, and comparing them with the posterior over hidden causes inferred by continuous data. However, the agent may fail to infer the causes of movement if the targets are constantly in motion, because the reduced models cannot be updated sufficiently quickly. A similar problem arises in the second example: in this case, a discrete model is needed to infer the correct sequence of actions needed to grasp the object, i.e., reach and then grasp. Again, the agent would fail the task if the object is moving too fast. Further, the object might also move during the grasping phase, so the agent has to carefully balance both actions depending on the object position at each continuous instant. The period of evidence accumulation actually poses a limit on how frequently the model can change the priors that must be realized or inferred.

In such cases, the agent's set of hypotheses must evolve in accordance with the external world. This raises a key question: how can agents perform discrete inference and planning when the properties of the environment are constantly changing? Here, we address this issue by treating discrete outcomes directly as models of trajectories and not fixed set points. Such trajectories can be seen as hypothetical evolutions of the external world, which can be used either for inferring the actual evolution (e.g., related to the intentions of the self or other agents) or for imposing a particular evolution *to* the world (i.e., achieving dynamic goals by acting).

Formally, we want to construct a flexible discrete representation of the environment modeled by the following system of stochastic differential equations:

$$\begin{aligned} \tilde{\boldsymbol{y}} &= \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}) + \tilde{\boldsymbol{w}}_y \\ \mathcal{D}\tilde{\boldsymbol{x}} &= \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}) + \tilde{\boldsymbol{w}}_x \end{aligned} \tag{1}$$

Here, $\tilde{\boldsymbol{y}}$, $\tilde{\boldsymbol{g}}$, $\tilde{\boldsymbol{x}}$, $\tilde{\boldsymbol{f}}$ are respectively the continuous observations, likelihood function, continuous hidden states, and dynamics function, while $\tilde{\boldsymbol{w}}_y$ and $\tilde{\boldsymbol{w}}_x$ are noise terms sampled from Gaussian distributions, all encoded in generalized coordinates of motion – see Section 4.1 for more details.

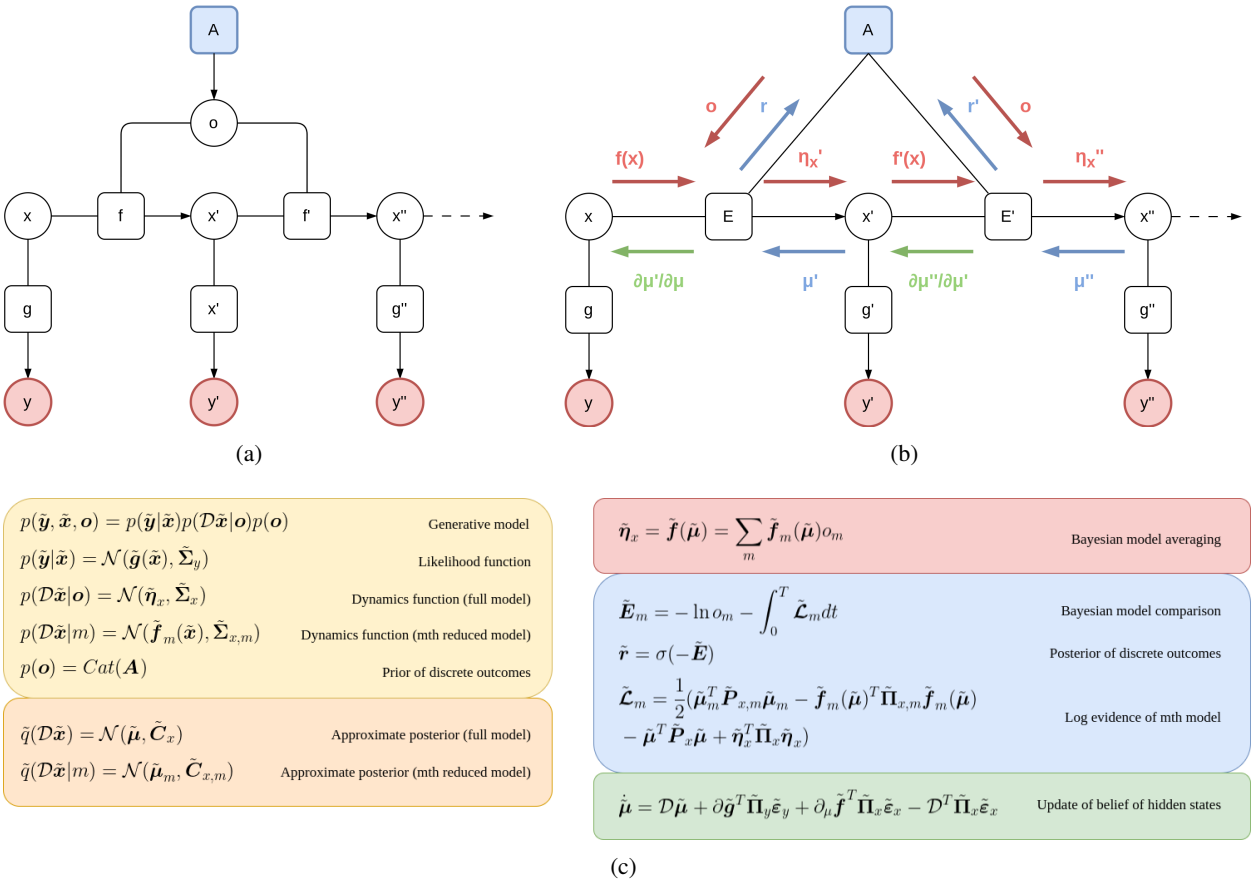


Figure 1: (a) Factor graph of a dynamic hybrid model. Discrete outcomes $\boldsymbol{o}$ encoding a dynamic set of hypotheses are generated by a prior $\boldsymbol{A}$. A potential evolution of the hidden states $\tilde{\boldsymbol{x}}$ is generated by a dynamics function $\tilde{\boldsymbol{f}}_m$, related to the mth dynamic hypothesis. In turn, the hidden states generate sensory predictions via the likelihood function $\tilde{\boldsymbol{g}}$, which are then compared with the corresponding sensations $\tilde{\boldsymbol{y}}$. (b) Graph exemplifying the exchange of messages within a dynamic hybrid model. The full prior $\tilde{\boldsymbol{\eta}}_x$ represents an average trajectory that depends on the potential trajectories and the discrete outcomes. In turn, the gradient $\partial_\mu \tilde{\boldsymbol{\mu}}$ indicates how the average trajectory is mapped back into previous orders. Last, the posterior $\tilde{\boldsymbol{r}}$ conveys the most likely agent's hypothesis that could have generated the current trajectory, wherein each temporal order refines the estimation. This is done by computing the free energy of every reduced model and temporal order, i.e., $\boldsymbol{E}$, $\boldsymbol{E}'$, and so on. (c) Generative model, approximate posteriors, Bayesian model averaging, Bayesian model comparison, and update rule of the beliefs.

We make use of the following hybrid generative model, displayed in Figure 1a:

$$\begin{aligned} p(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{x}}, \boldsymbol{o}) &= p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}})p(\mathcal{D}\tilde{\boldsymbol{x}}|\boldsymbol{o})p(\boldsymbol{o}) \\ p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{\Sigma}}_y) \\ p(\mathcal{D}\tilde{\boldsymbol{x}}|\boldsymbol{o}) &= \mathcal{N}(\tilde{\boldsymbol{\eta}}_x, \tilde{\boldsymbol{\Sigma}}_x) \\ p(\boldsymbol{o}) &= Cat(\boldsymbol{A}) \end{aligned} \tag{2}$$

where $\boldsymbol{A}$ is a prior. In particular, the likelihood and dynamics are encoded with Gaussian probability distributions. Instead, discrete outcomes $\boldsymbol{o}$ are generated by a categorical distribution, wherein each element represents the probability that a specific continuous trajectory will be perceived, i.e., $\boldsymbol{o} = [\boldsymbol{o}_1, \ldots, \boldsymbol{o}_M]$, where M is the number of dynamic hypotheses. Further, we consider the distribution of the dynamics $p(\mathcal{D}\tilde{\boldsymbol{x}}|\boldsymbol{o})$ as a full model over the set of hypotheses, where $\tilde{\boldsymbol{\eta}}_x$ is the generalized prior over the trajectory. We can infer the most likely discrete outcome under the prior $\boldsymbol{A}$ if

we assume that the agent maintains M reduced models:

$$p(\mathcal{D}\tilde{\boldsymbol{x}}|m) = \mathcal{N}(\tilde{\boldsymbol{f}}_m(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{\Sigma}}_{x,m}) \tag{3}$$

with the same likelihood as the full model but differing in their prior. Here, $\tilde{\boldsymbol{f}}_m(\tilde{\boldsymbol{x}})$ is the dynamics function of the mth reduced model. Following the variational approach described in Section 4.3, we define the corresponding approximate reduced and full models:

$$\begin{aligned} \tilde{q}(\mathcal{D}\tilde{\boldsymbol{x}}|m) &= \mathcal{N}(\tilde{\boldsymbol{\mu}}_m, \tilde{\boldsymbol{C}}_{x,m}) \\ \tilde{q}(\mathcal{D}\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{C}}_x) \end{aligned} \tag{4}$$

The implications of Equations 3 and 4 can be better understood from Figure 1b, which highlights three different kinds of message passing between discrete and continuous representations. Top-down messages are computed through a Bayesian model average of the prior trajectories $\tilde{\boldsymbol{f}}_m(\tilde{\boldsymbol{\mu}})$ of each outcome model and the corresponding probabilities:

$$\tilde{\boldsymbol{\eta}}_x = \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{\mu}}) = \sum_m \tilde{\boldsymbol{f}}_m(\tilde{\boldsymbol{\mu}}) o_m \tag{5}$$

As a consequence, the agent does not maintain fixed priors over the environment, but dynamic priors that depend on potential trajectories and evolve at each continuous time step.

Concerning the inference, each reduced model is scored as in conventional hybrid models, but now the posterior over the current trajectories is compared with the trajectory generated by the dynamics functions of the hidden states:

$$\tilde{\boldsymbol{\mathcal{L}}}_m = \frac{1}{2}(\tilde{\boldsymbol{\mu}}_m^T \tilde{\boldsymbol{P}}_{x,m} \tilde{\boldsymbol{\mu}}_m - \tilde{\boldsymbol{f}}_m(\tilde{\boldsymbol{\mu}})^T \tilde{\boldsymbol{\Pi}}_{x,m} \tilde{\boldsymbol{f}}_m(\tilde{\boldsymbol{\mu}}) - \tilde{\boldsymbol{\mu}}^T \tilde{\boldsymbol{P}}_x \tilde{\boldsymbol{\mu}} + \tilde{\boldsymbol{\eta}}_x^T \tilde{\boldsymbol{\Pi}}_x \tilde{\boldsymbol{\eta}}_x) \tag{6}$$

expressed in terms of posterior and prior precisions $\tilde{\boldsymbol{P}}_x$ and $\tilde{\boldsymbol{\Pi}}_x$. Here, the posterior mean and precision of each reduced model are:

$$\begin{aligned} \tilde{\boldsymbol{\mu}}_m &= \tilde{\boldsymbol{C}}_{x,m}(\tilde{\boldsymbol{P}}_x \tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{\Pi}}_x \tilde{\boldsymbol{\eta}}_x + \tilde{\boldsymbol{\Pi}}_{x,m} \tilde{\boldsymbol{f}}_m(\tilde{\boldsymbol{\mu}})) \\ \tilde{\boldsymbol{P}}_{x,m} &= \tilde{\boldsymbol{P}}_x - \tilde{\boldsymbol{\Pi}}_x + \tilde{\boldsymbol{\Pi}}_{x,m} \end{aligned} \tag{7}$$

At this point, for each temporal order and model, a Bayesian model comparison is performed between the prior surprise and the log evidence accumulated over a time T. Finally, a softmax function is applied to get a proper probability:

$$\tilde{\boldsymbol{r}} = \sigma(-\tilde{\boldsymbol{E}}) = \sigma(\ln \boldsymbol{o} + \int_0^T \tilde{\boldsymbol{\mathcal{L}}}_m dt) \tag{8}$$

In summary, what Equation 5 means is that the agent has some dynamic hypotheses over the possible causes that could have generated its perceptions. It hence makes a guess by combining the related potential trajectories. The other side of the flow works similarly: the agent accumulates, for a certain amount of time, a score – i.e., the log evidence $\tilde{\boldsymbol{\mathcal{L}}}_m$ – of such dynamic hypotheses depending on continuous data, that is, it infers the most likely model that could have produced the perceived trajectory. The posterior $\tilde{\boldsymbol{r}}$ over discrete outcomes – expressed by Equation 8 – defines the agent's final guess after observing the evidence, where each temporal order further refines the estimation. Since the potential trajectories are continuously generated, a model can be inferred that corresponds to a dynamic path for the whole period of evidence accumulation.

In relation to the backward pass – highlighted in green in Figure 1a – the prediction error over dynamics between the actual trajectory and the full prior, i.e., $\tilde{\boldsymbol{\varepsilon}}_x = \mathcal{D}\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{\eta}}_x$, is backpropagated via the gradient of the dynamics functions with respect to the hidden states:

$$\dot{\tilde{\boldsymbol{\mu}}} \propto \frac{\partial \tilde{\boldsymbol{f}}}{\partial \tilde{\boldsymbol{\mu}}}^T \tilde{\boldsymbol{\Pi}}_x \tilde{\boldsymbol{\varepsilon}}_x \tag{9}$$

This gradient encodes how an average 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.

This architecture has a key implication related to the tasks mentioned above. If o_m is set to 1 and all the others to 0, the hidden states will be subject only to the mth trajectory. Due to the dualism between action and perception, if multiple outcome models 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, discrete outcomes not only represent the agent's beliefs about the environment but also its intentions, now encoded with dynamic representations. As a result, a specific reaching movement can be found by generating and comparing dynamic hypotheses depending on the trajectories of the hand and every target, all inferred from continuous data. Additionally, grasping a moving object is possible by linking the prior to a discrete model, i.e., $\boldsymbol{o}_\tau = \boldsymbol{A}\boldsymbol{s}_\tau$, where $\boldsymbol{A}$ is the likelihood matrix and $\boldsymbol{s}_\tau$ are the discrete hidden states at time τ – see Section 4.2. This process unfolds into the following steps: (i) the discrete model predicts an outcome corresponding to the reaching action; (ii) the discrete outcome steers the lower-level dynamics by generating a specific trajectory; (iii) this 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 discrete outcome based on the accumulated evidence – i.e., it signals whether the hypothesis has been realized – and the discrete model eventually predicts the next (grasping) action. Because a discrete model can maintain a high-level intention for the whole period of evidence accumulation, a combined reaching-grasping action naturally emerges via inference – as we demonstrate in the simulations.

2.2 Two examples of dynamic inference

Ultimately, the choice of hybrid model depends on the task considered, with both types often necessary in realistic settings. Here, we briefly describe two examples where a dynamic hybrid model is fundamental. In the first one, the agent infers its intention 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 its goal is to infer which one it is following. The dynamic nature of the task can be tackled by defining potential trajectories for the objects of interest. Figure 2 shows a graphical representation of the hybrid generative model used by the agent. This is a basic inference task that does not involve high-level planning, thus we do not consider a complete discrete model. The agent controls an arm with a single Degree of Freedom (DoF). Its sensory modalities are: (i) a proprioceptive observation y_p for the arm's joint angle; (ii) visual observations $\tilde{\boldsymbol{y}}_v$ encoding the Cartesian positions and velocities of the hand and targets.

Since encoding Cartesian positions is not strictly necessary for this task, we express hidden states $\boldsymbol{x}$ in joint angles only, which thus generate and integrate both sensory modalities. The hidden states encode information about all the three elements of the environment:

$$\boldsymbol{x} = \begin{bmatrix} x_a & x_{t1} & x_{t2} \end{bmatrix} \tag{10}$$

while the successive order $\boldsymbol{x}'$ expresses their angular velocities. Proprioceptive predictions – needed to infer the kinematic configuration of the agent's body – are generated by a function g_p that simply extracts the arm's joint angle from the belief $\tilde{\boldsymbol{\mu}}$ of the hidden states:

$$g_p(\boldsymbol{\mu}) = \mu_a \tag{11}$$

Instead, visual predictions are generated by a function $\tilde{\boldsymbol{g}}_v$ that computes the Cartesian positions and velocities of every component of the hidden states:

$$\begin{aligned} \tilde{\boldsymbol{g}}_v(\tilde{\boldsymbol{\mu}}) &= \begin{bmatrix} \boldsymbol{T}(\mu_a) & \boldsymbol{T}(\mu_{t1}) & \boldsymbol{T}(\mu_{t2}) \\ \boldsymbol{J}(\mu_a)\mu'_a & \boldsymbol{J}(\mu_{t1})\mu'_{t1} & \boldsymbol{J}(\mu_{t2})\mu'_{t2} \end{bmatrix} \\ \boldsymbol{T}(\theta) &= \begin{bmatrix} lc_\theta & ls_\theta \end{bmatrix} \\ \boldsymbol{J}(\theta) &= \begin{bmatrix} -ls_\theta & lc_\theta \end{bmatrix} \end{aligned} \tag{12}$$

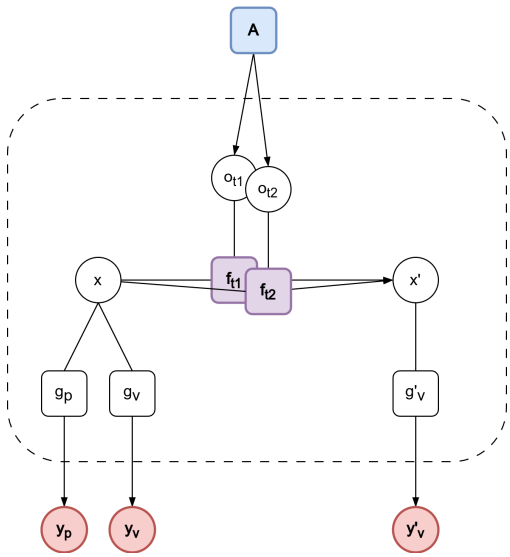


Figure 2: Factor graph for the trajectory inference task. Hidden states $\tilde{\boldsymbol{x}}$ for the arm and the targets – encoded in joint angles – generate proprioceptive and visual predictions via likelihood functions g_p and $\tilde{\boldsymbol{g}}_v$, respectively. Through Bayesian model comparison, the most likely agent's intention – encoded as discrete outcomes $\boldsymbol{o}$ – is dynamically inferred.

where $\boldsymbol{T}$ is the forward kinematics, $\boldsymbol{J}$ its Jacobian, l is the arm length, while s_θ and c_θ are the sine and cosine of the angle.

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

$$\varepsilon_p = y_p - g_p(\boldsymbol{\mu}) \qquad\qquad \tilde{\boldsymbol{\varepsilon}}_v = \tilde{\boldsymbol{y}}_v - \tilde{\boldsymbol{g}}_v(\tilde{\boldsymbol{\mu}}) \tag{13}$$

Discrete outcomes $\boldsymbol{o}$ – which we let here depend on a uniform prior – have only two elements, corresponding to the agent's intentions to follow the first or the second target:

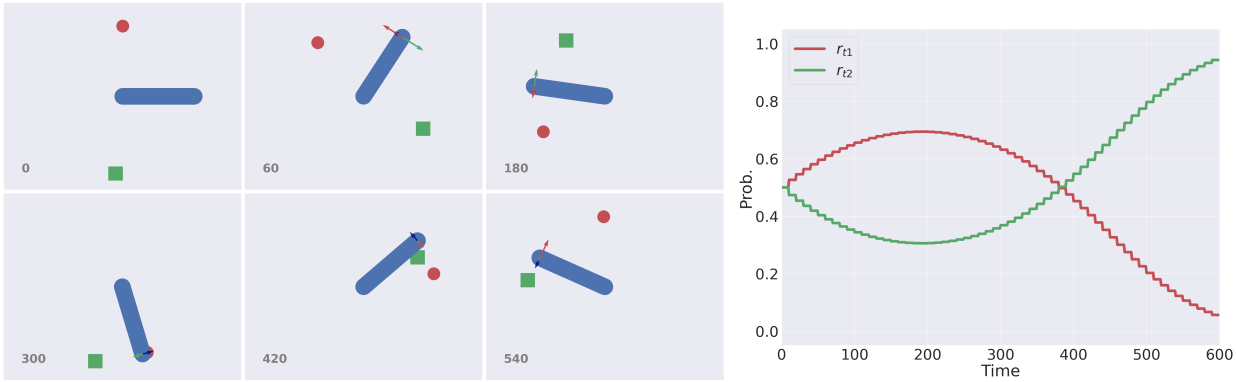
$$\boldsymbol{o} = \begin{bmatrix} o_{t1} & o_{t2} \end{bmatrix} \tag{14}$$

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 intentional future states in which the first component (corresponding to the arm's joint angle) is equal to the inferred angle of the objects, i.e., $\boldsymbol{i}_{t1}(\boldsymbol{\mu}) = [\mu_{t1}, \mu_{t1}, \mu_{t2}]$ and $\boldsymbol{i}_{t2}(\boldsymbol{\mu}) = [\mu_{t2}, \mu_{t1}, \mu_{t2}]$. Next, we define a dynamics function for each reaching intention:

$$\begin{aligned} \boldsymbol{f}_{t1}(\boldsymbol{\mu}) &= \lambda(\boldsymbol{i}_{t1}(\boldsymbol{\mu}) - \boldsymbol{\mu}) \\ \boldsymbol{f}_{t2}(\boldsymbol{\mu}) &= \lambda(\boldsymbol{i}_{t2}(\boldsymbol{\mu}) - \boldsymbol{\mu}) \end{aligned} \tag{15}$$

where λ is an attractor gain or precision. Such functions encode potential trajectories that the agent may observe, and correspond to attractors pointing toward a target – that is, the agent thinks that its arm will be pulled toward the inferred target angle. The dynamic prior $\boldsymbol{\eta}'_x$ is found by weighting the potential trajectories with their corresponding probabilities that they will be perceived, i.e., o_{t1} and o_{t2}. This is done through Equation 5, which is then used to compute a dynamics prediction error $\boldsymbol{\varepsilon}_x$:

$$\begin{aligned} \boldsymbol{\eta}'_x &= \boldsymbol{f}(\boldsymbol{\mu}) = o_{t1}\boldsymbol{f}_{t1}(\boldsymbol{\mu}) + o_{t2}\boldsymbol{f}_{t2}(\boldsymbol{\mu}) \\ \boldsymbol{\varepsilon}_x &= \boldsymbol{\mu}' - \boldsymbol{\eta}'_x \end{aligned} \tag{16}$$



(a) (b)

Figure 3: A simple inference task in which the agent (in blue) has to infer which one of two moving targets (a red ball and a green square) is following. See also [32] for a similar inference task in the Cartesian domain. (a) A sample sequence of time frames. The arm and the two targets move around the origin with constant but different velocities. The actual trajectory of the arm is represented with a blue arrow, while the potential trajectories of the two targets are displayed with red and green arrows. (b) Posterior over discrete outcomes, encoding the probability of reaching the red ball (r_{t1}) and the green square (r_{t2}), over time. The stepped behavior of the trajectories are due to the window T of continuous evidence accumulation of Equation 8, which was set to 10 continuous time steps.

The belief of the hidden states is updated by minimizing the free energy:

$$\dot{\tilde{\boldsymbol{\mu}}} = \begin{bmatrix} \boldsymbol{\mu}' + \partial g_p^T \pi_p \varepsilon_p + \partial \boldsymbol{g}_v^T \pi_v \boldsymbol{\varepsilon}_v + \partial \boldsymbol{f}^T \pi_x \boldsymbol{\varepsilon}_x \\ \partial \boldsymbol{g}_v'^T \pi_p' \boldsymbol{\varepsilon}_v' - \pi_x \boldsymbol{\varepsilon}_x \end{bmatrix} \tag{17}$$

and using gradient descent. On the other hand, trajectory inference is done via Equations 6 and 8, where the posteriors over the reduced models are computed by Equation 7.

Figure 3 illustrates a simulation of the presented task. The arm and the two targets move at constant angular velocities. However, the red circle and the green square move slower and faster than the arm, respectively (see Figure 3a). As shown in Figure 3b, both discrete outcomes are initialized with equal probabilities, but the first outcome slowly increases as the arm approaches the first target. Then, it reaches a maximum, after which it starts to slowly decrease until the second outcome takes over. As explained before, the log evidence L_m of the mth reduced model scores how well the potential angular velocity of a target can explain the actual arm's angular velocity (whose direction and magnitude are shown by the arrows in Figure 3a). The posterior $\boldsymbol{r}$ over discrete outcomes accumulate the evidence for the whole period in which the objects are moving, inferring the correct probability for each intention. Note that with an uninformative prior, Bayesian model comparison reduces to evidence accumulation. Overall, this simple task demonstrates that the model can discriminate between causes that generate continuous trajectories.

2.2.2 Object grasping

Grasping a moving object requires more complex computations than trajectory inference. In this case, the agent's body is an 8-DoF arm where the last four joints correspond to the fingers. We use three observations: (i) proprioceptive observations $\boldsymbol{y}_p = [y_{p,1}, \ldots, y_{p,8}]$ encoding the joint angles of the arm; (ii) visual observations $\tilde{\boldsymbol{y}}_v = [\tilde{\boldsymbol{y}}_{v,a}, \tilde{\boldsymbol{y}}_{v,o}]$ encoding the Cartesian positions and velocities of the hand and object; (iii) a discrete tactile outcome $\boldsymbol{o}_t$ signaling whether the hand touches the object. We construct a discrete model with hidden states $\boldsymbol{s}$ encoding: (i) whether or not the hand is at the object 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. The transitions encoded in the transition matrix $\boldsymbol{B}$ correspond to the three steps of the task, plus a *stay* action: (i) open the hand; (ii) reach the target; (iii) close the hand. The transition matrix is defined such that the object can be grasped only when both hand and object are in the same position.

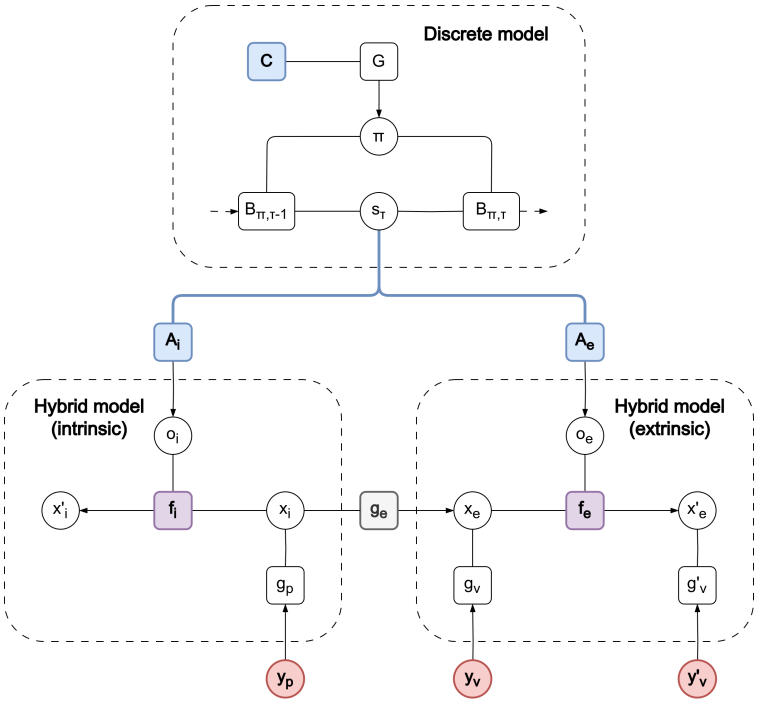


Figure 4: Factor graph for the object grasping task. Discrete hidden states $\boldsymbol{s}$ at time τ compute in parallel two predictions of discrete outcomes ($\boldsymbol{o}_i$ and $\boldsymbol{o}_e$) for the intrinsic and extrinsic modalities via likelihood matrices $\boldsymbol{A}_i$ and $\boldsymbol{A}_e$. In turn, the discrete outcomes bias the continuous hidden states below ($\tilde{\boldsymbol{x}}_i$ and $\tilde{\boldsymbol{x}}_e$), thus generating a particular trajectory. The two modalities are linked by a generative model $\boldsymbol{g}_e$ that achieves kinematic inversion via inference, and they generate proprioceptive and visual predictions, respectively.

We then define two dynamic hybrid models expressing intrinsic and extrinsic modalities (whose factor graphs are illustrated in Figure 4). Their generalized hidden states ($\tilde{\boldsymbol{x}}_i$ and $\tilde{\boldsymbol{x}}_e$) respectively encode information in intrinsic (e.g., joint angles and angular velocities) and extrinsic (e.g., Cartesian positions and velocities) coordinates. Moreover, both modalities comprise components corresponding to the arm and object, reflecting the decomposition of visual observations:

$$\tilde{\boldsymbol{x}}_i = \begin{bmatrix} \tilde{\boldsymbol{x}}_{i,a} & \tilde{\boldsymbol{x}}_{i,o} \end{bmatrix} \qquad \tilde{\boldsymbol{x}}_e = \begin{bmatrix} \tilde{\boldsymbol{x}}_{e,a} & \tilde{\boldsymbol{x}}_{e,o} \end{bmatrix} \tag{18}$$

The 0th-order extrinsic hidden states $\boldsymbol{x}_{e,a}$ and $\boldsymbol{x}_{e,o}$ encode the positions of the hand and object, respectively. About the 0th-order intrinsic hidden states, while the interpretation of $\boldsymbol{x}_{i,a}$ is intuitive (it just represents the actual body configuration related to the hand position), the other component can be seen as encoding a potential body configuration related to the object position [33]. In short, it may be considered as a state encoding a particular interaction with the object, depending on its affordances. The two modalities are linked via the 0th-order hidden states, by a likelihood function $\boldsymbol{g}_e$ that performs forward kinematics:

$$\boldsymbol{g}_e(\boldsymbol{x}_i) = \begin{bmatrix} \boldsymbol{T}(\boldsymbol{x}_{i,a}) & \boldsymbol{T}(\boldsymbol{x}_{i,o})) \end{bmatrix} \tag{19}$$

Since we are not interested in the fingers' positions during the reaching task, these functions only compute the position of the middle of the two fingers. Maintaining a belief $\boldsymbol{\mu}_e$ in extrinsic coordinates greatly simplifies the computations because it permits expressing the dynamics without worrying about the inverse mapping to the joint angles, which is achieved by inverting the kinematic model via inference – for more details, see [34].

The discrete hidden states $\boldsymbol{s}$ simultaneously generate predictions for the hidden causes of intrinsic and extrinsic modalities through likelihood matrices $\boldsymbol{A}_i$ and $\boldsymbol{A}_e$. With this architecture, we can separately express grasping and

reaching intentions in the two modalities:

$$\begin{aligned} \boldsymbol{o}_i &= \begin{bmatrix} o_{i,open} & o_{i,close} \end{bmatrix} \\ \boldsymbol{o}_e &= \begin{bmatrix} o_{e,stay} & o_{e,reach} \end{bmatrix} \end{aligned} \tag{20}$$

Nonetheless, more advanced movements – such as grasping an object with a specific wrist rotation – can be achieved by also expressing reaching intentions in the intrinsic modality [33]. Last, a matrix $\boldsymbol{A}_t$ returns a prediction for the tactile observation $\boldsymbol{o}_t$.

The discrete outcomes are linked to potential trajectories defining grasping and reaching intentions in the continuous domain. For the intrinsic modality:

$$\begin{aligned} \boldsymbol{i}_{i,open}(\boldsymbol{\mu}_i) &= \begin{bmatrix} \boldsymbol{b}_{i,open} & \boldsymbol{\mu}_{i,o} \end{bmatrix} \\ \boldsymbol{i}_{i,close}(\boldsymbol{\mu}_i) &= \begin{bmatrix} \boldsymbol{b}_{i,close} & \boldsymbol{\mu}_{i,o} \end{bmatrix} \end{aligned} \tag{21}$$

where $\boldsymbol{b}_{i,open}$ and $\boldsymbol{b}_{i,close}$ are the closed/open configurations. Correspondingly, for the extrinsic modality:

$$\begin{aligned} \boldsymbol{i}_{e,stay}(\boldsymbol{\mu}_e) &= \boldsymbol{\mu}_e \\ \boldsymbol{i}_{e,reach}(\boldsymbol{\mu}_e) &= \begin{bmatrix} \boldsymbol{\mu}_{e,o} & \boldsymbol{\mu}_{e,o} \end{bmatrix} \end{aligned} \tag{22}$$

Here, $\boldsymbol{i}_{e,stay}$ is an identity mapping and can be seen as an intention to maintain the current state of the world. Finally, we associate a dynamics function for each of these future states, i.e., $\boldsymbol{f}_{i,open}$, $\boldsymbol{f}_{i,close}$, $\boldsymbol{f}_{e,stay}$, and $\boldsymbol{f}_{e,reach}$.

The update of the hidden states follows the derivations of the previous example:

$$\begin{aligned} \dot{\tilde{\boldsymbol{\mu}}}_i &= \begin{bmatrix} \boldsymbol{\mu}_i' + \partial \boldsymbol{g}_p^T \pi_p \boldsymbol{\varepsilon}_p + \partial \boldsymbol{f}_i^T \pi_{x,i} \boldsymbol{\varepsilon}_{x,i} \\ -\pi_{x,i} \boldsymbol{\varepsilon}_{x,i} \end{bmatrix} \\ \dot{\tilde{\boldsymbol{\mu}}}_e &= \begin{bmatrix} \boldsymbol{\mu}_e' + \partial \boldsymbol{g}_e^T \pi_e \boldsymbol{\varepsilon}_e + \partial \boldsymbol{g}_v^T \pi_v \boldsymbol{\varepsilon}_v + \partial \boldsymbol{f}_e^T \pi_{x,e} \boldsymbol{\varepsilon}_{x,e} \\ \partial \boldsymbol{g}_v'^T \pi_v' \boldsymbol{\varepsilon}_v' - \pi_{x,e} \boldsymbol{\varepsilon}_{x,e} \end{bmatrix} \end{aligned} \tag{23}$$

where $\boldsymbol{g}_p$ and $\tilde{\boldsymbol{g}}_v$ are simple identity mappings, while $\boldsymbol{f}_i$ and $\boldsymbol{f}_e$ are the functions computing a Bayesian model average of intrinsic and extrinsic trajectories. As concerns the posteriors over discrete outcomes at discrete time τ, they are updated by combining the agent's surprise with continuous evidence:

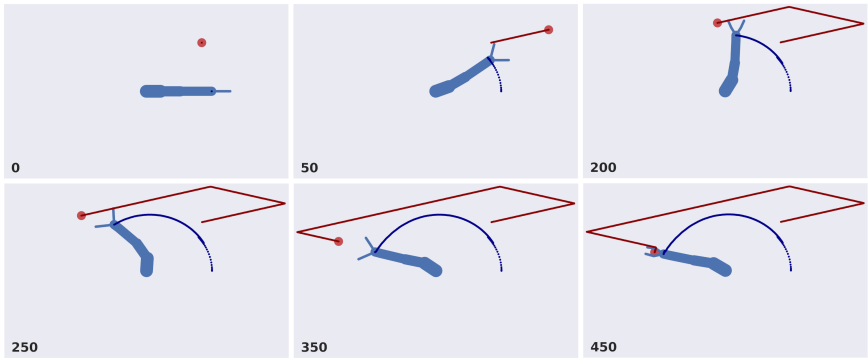
$$\begin{aligned} \boldsymbol{r}_{i,\tau} &= \sigma(\ln \boldsymbol{A}_i \boldsymbol{s}_\tau + \int_0^T \boldsymbol{\mathcal{L}}_i dt) \\ \boldsymbol{r}_{e,\tau} &= \sigma(\ln \boldsymbol{A}_e \boldsymbol{s}_\tau + \int_0^T \boldsymbol{\mathcal{L}}_e dt) \end{aligned} \tag{24}$$

These messages are finally used by the discrete model, along with the tactile observation, to infer the most probable discrete hidden state at time τ and conditioned over policy π:

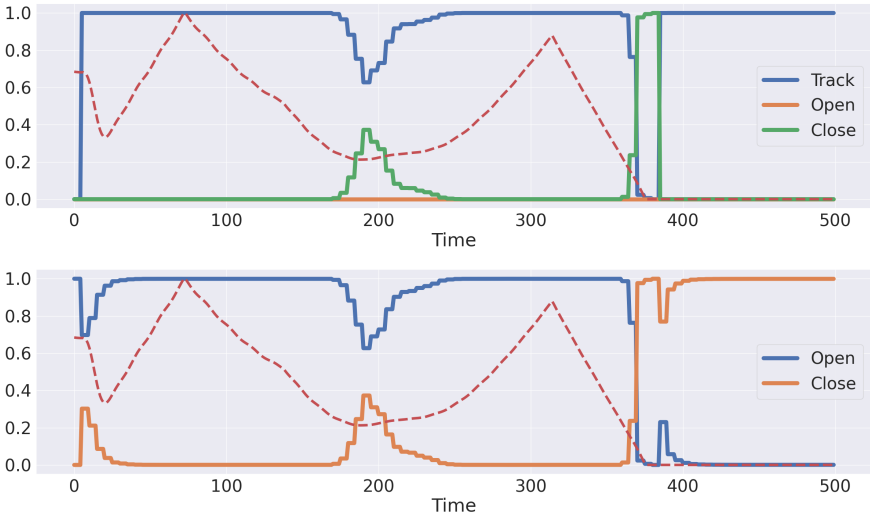
$$\boldsymbol{s}_{\pi,\tau} = \sigma(\ln \boldsymbol{B}_{\pi,\tau-1} \boldsymbol{s}_{\tau-1} + \ln \boldsymbol{A}_i^T \boldsymbol{r}_{i,\tau} + \ln \boldsymbol{A}_e^T \boldsymbol{r}_{e,\tau} + \ln \boldsymbol{A}_t^T \boldsymbol{o}_t) \tag{25}$$

Note that the evidence accumulation is performed within the two dynamic hybrid models, so that the discrete model only needs to combine the two bottom-up messages, infer the most probable state, perform planning operations, and finally send predicted intentions back to the low levels, thus synchronizing their behavior.

Since the two hybrid models are now controlled by a discrete model, the agent can make a new plan if a grasping action fails because the object is moving too fast. This behavior can be seen in Figure 5b, which displays the trajectories of discrete hidden states $\boldsymbol{s}$ and intrinsic outcomes $\boldsymbol{o}_i$. As soon as the hand approaches the object (at t=200), the agent makes a first grasping attempt, resulting in a gradual decrease of the hand-at-object state and a gradual increase of the



(a)



(b)

Figure 5: Object grasping with a 4-DoF arm and a 4-DoF hand with two fingers. (a) A sequence of time frames showing the arm (in blue) grasping the moving target (red ball). Trajectories of hand and object are shown, too. (b) Discrete hidden states $\boldsymbol{s}$ (top panel) and posterior over intrinsic discrete outcomes $\boldsymbol{r}$ (bottom panel), over time. Both panels also show the hand-object distance (dashed red line), normalized between 0 and 1.

closed-hand state. These states generate discrete outcomes in both modalities so that the *open* and *close* trajectories change accordingly. However, since the object is still too far from the hand, the agent rapidly turns back to tracking. At about t=400, the object is close enough to the hand, resulting in a much steeper change in the discrete hidden states and intrinsic outcomes. Eventually, the object is correctly grasped.

3 Discussion

We presented a method that treats two fundamental problems – intention inference and motor control – under an active inference hybrid framework that relies on dynamic representations. The method leverages Bayesian model reduction [17], a powerful technique that, among its wide usages, allows shrinking continuous signals into a small set of hypotheses associated with discrete categories. In traditional active inference models, discrete outcomes are combined with static continuous priors to find the most likely hidden states based on the current evidence. This approach can be used, e.g., to reach fixed positions in sequence [1] or simulate pictographic reading [25]. However, using static assumptions limits applicability in real-world scenarios where goals and trajectories unfold dynamically. Our approach addresses this limitation by maintaining a set of dynamic hypotheses, that is, discrete outcomes related to potential

trajectories. From the point of view of the continuous model, the discrete outcomes directly act as hidden causes, and a Bayesian model reduction is performed across multiple temporal orders of the hidden states.

This method permits drawing useful parallelisms on action understanding and policy evaluation. Specifically, top-down messages can be interpreted as the agent's average intended trajectory (what the agent aims to realize), while bottom-up messages reflect the most likely intention that could have generated the observed movements. With this approach, evidence can be dynamically accumulated and associated with specific intentions. This allows the agent to constantly evaluate which potential evolution of the external world best explains its sensory observations, and act appropriately. This can be useful in every kind of context that presents dynamic elements, e.g., when interacting with moving objects.

An interesting fact regards the inference of self-generated movements. As also shown in [35], intention understanding in dynamic environments and embodied decision-making are two emergent properties of the proposed architecture. Although it may seem counterintuitive, inferring a decision from one's own perceptions and movements is actually a key aspect of biological behavior that helps to reinforce and stabilize the decision taken [36, 37], ensuring that valuable opportunities are not missed in ecological settings. This mechanism has received little attention in artificial intelligence and robotics, but it could support advances in the next generation of intelligent agents. Finally, the presented model can be scaled up to control complex kinematic structures, synchronizing the behavior of multiple limbs and performing complex operations, e.g., for tool use [34, 38].

The interaction between inference and action becomes even clearer when we compare the precisions $\tilde{\mathbf{\Pi}}_{x,m}$ of the reduced dynamics functions of Equation 6 and the sensory precisions $\tilde{\mathbf{\Pi}}_{y}$ – c.f., attentional gain or selection. In predictive coding, sensory precisions both encode the agent's confidence about a sensory modality (e.g., when visual input is clear and unambiguous, visual precision is high), and attention (e.g., in a noisy environment, auditory precision may be low, so the agent shifts reliance to vision) [39]. In our case, a high precision assigned to a potential trajectory means it is a good option for minimizing the free energy or, more clearly, for inferring the current agent's trajectory or realizing its goal. Conversely, a low precision implies that the trajectory is not useful for the task – e.g., a grasping action when the object is out of reach – or that it cannot be helpful to understand the situation – e.g., ignoring a target unrelated to the motion of another agent. In fact, attention is closely related to intention, in the sense that both are selected on the basis of relative precision or confidence. Attentional selection relates to state estimation, while intentional selection relates to action and execution. When considering learning of such precisions, two concurrent processes arise: a *fast* process (described here) that selects trajectories based on current goals and evidence; and a *slow* learning process that updates how useful each trajectory is in different contexts – in the same way that unreliable sensors get lower precision over time [40].

Overall, the model that we presented might be helpful to understand how dynamic hypotheses 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 5b, although we used elementary reaching and grasping actions, a smooth approach transition emerges, due to the Bayesian model reduction between potential trajectories. This novel behavior could be mapped into the continuous dynamics and called in a single action by the discrete model, producing a flexible and fluid behavior – as simulated in [33].

Future work may extend this approach in several directions. One limitation of this study is that we kept the precisions of the dynamics functions fixed. Learning the dynamics functions of the reduced models, along with their precisions, would be a promising direction of research to simulate task specialization and construction of richer behaviors from primitive movements. To address this issue, an alternative approach was implemented in [41], which used multiple continuous generative models followed by a "switcher". This approach has some analogies with the model presented here in that the agent (a student bird) could maintain models of how teachers generate a possible evolution of sensory signals, and through the switcher it could perform online model selection to infer the most likely causes. Another study proposed a hybrid active inference model based on recurrent switching linear dynamical systems (rSLDS) [42]. This allowed the agent to learn discrete representations of continuous dynamics via piecewise linear decomposition, useful to solve complex tasks such as the Mountain Car problem.

Another limitation is the use of just two temporal orders, whereas trajectory inference with increasing temporal orders might be helpful to develop more precise robots or, for neuroscience, to decode intentions from gaze or other highly dynamic cues. Finally, the model could be used for multi-agent interaction, where an agent has to observe and imitate the intentions of other agents, either in collaborative or competitive settings.

4 Methods

4.1 Active inference in continuous time

Active inference is a brain-inspired theory that proposes a unifying paradigm to understand the behavior of all living organisms. It is based upon the *free energy principle* which states that, in order to survive, every creature must actively minimize the *free energy* [1, 2, 3, 4], which is the opposite of what is known in the machine learning community as *evidence lower bound* (or ELBO). The theory of active inference makes the same assumptions of predictive coding, i.e., that organisms perceive their environment through an internal generative model, inferring how external causes (of the real generative process) lead to the observed sensory signals [4, 2]. Regarding the continuous domain, this procedure is broken down into probability distributions comprising hidden causes $\tilde{\boldsymbol{v}}$, hidden states $\tilde{\boldsymbol{x}}$, and sensory signals $\tilde{\boldsymbol{y}}$:

$$p(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}, \tilde{\boldsymbol{y}}) = p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}})p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}})p(\tilde{\boldsymbol{v}}) \tag{26}$$

These distributions are approximated through Gaussian functions:

$$\begin{aligned} p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{\Sigma}}_y) \\ p(\mathcal{D}\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}}) &= \mathcal{N}(\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}), \tilde{\boldsymbol{\Pi}}_x) \\ p(\tilde{\boldsymbol{v}}|\boldsymbol{\eta}_v) &= \mathcal{N}(\boldsymbol{\eta}_v, \tilde{\boldsymbol{\Pi}}_v) \end{aligned} \tag{27}$$

which respectively correspond to: (i) the likelihood function $\tilde{\boldsymbol{g}}$, encoding how hidden states generate sensory signals; (ii) the dynamics function $\tilde{\boldsymbol{f}}$, specifying the evolution of the hidden states; (iii) the prior $\boldsymbol{\eta}_v$ over the hidden causes, whose role is to define targets for goal-directed behavior and to link hierarchical levels. Further, $\tilde{\boldsymbol{\Sigma}}_y$, $\tilde{\boldsymbol{\Sigma}}_x$, and $\tilde{\boldsymbol{\Sigma}}_v$ are the variances of sensory signals, hidden states, and hidden causes. Note that the symbol $\sim$ denotes a variable in generalized coordinates of motion, 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 $\mathcal{D}\tilde{\boldsymbol{x}} = [\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}''' \dots]$. These functions are related to the following nonlinear stochastic equations that describe the evolution of the environment:

$$\begin{aligned} \tilde{\boldsymbol{y}} &= \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}) + \boldsymbol{w}_y \\ \mathcal{D}\tilde{\boldsymbol{x}} &= \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}) + \boldsymbol{w}_x \end{aligned} \tag{28}$$

where $\tilde{\boldsymbol{w}}_y$ and $\tilde{\boldsymbol{w}}_x$ are noise terms sampled from Gaussian distributions.

In active inference, beliefs are 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 produces an action and makes the agent fulfill prior expectations. Belief dynamics can be seen as a combination of two components: (i) a model of how hidden states normally evolve based on past experience – e.g., expecting a moving ball to follow a straight path; (ii) a distorted vision of this dynamics that depends on the agent's desires – e.g., thinking that my hand will be pulled toward the ball will drive the movements of my hand toward my expectations, even if that force does not exist in reality.

The agent's goal is to compute the posterior $p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})$, whose direct computation is however 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 the following Gaussians:

$$\begin{aligned} q(\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{x}}; \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{C}}_x) \\ q(\tilde{\boldsymbol{v}}) &= \mathcal{N}(\tilde{\boldsymbol{v}}; \tilde{\boldsymbol{\nu}}, \tilde{\boldsymbol{C}}_v) \end{aligned} \tag{29}$$

where the parameters $\tilde{\boldsymbol{\mu}}$ and $\tilde{\boldsymbol{\nu}}$ are the beliefs of hidden states and hidden causes, respectively. Perceptual inference then becomes a minimization process, aiming to reduce the disparity between the approximated and actual posteriors. This can be framed in terms of a KL divergence – equivalent to the expectation, over the approximated posterior, of the discrepancy between the two log probability distributions:

$$DKL[q(\tilde{\boldsymbol{x}})||p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})] = \sum_x q(\tilde{\boldsymbol{x}}) \ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})} = \mathbb{E}\, q(\tilde{\boldsymbol{x}})[\ln q(\tilde{\boldsymbol{x}}) - \ln p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})] \tag{30}$$

As the denominator $p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})$ still depends 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 instead. Given that the KL divergence is non-negative (due to Jensen's inequality), the VFE serves as an upper bound on surprise. Consequently, minimizing the VFE enhances both the evidence and model fit:

$$\mathcal{F} = \mathbb{E}\, q(\tilde{\boldsymbol{x}}) \left[\ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})} \right] = \underset{q(\tilde{\boldsymbol{x}})}{\mathbb{E}} \left[\ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})} \right] - \ln p(\tilde{\boldsymbol{y}}) \tag{31}$$

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

$$\begin{aligned} \dot{\tilde{\boldsymbol{\mu}}} - \mathcal{D}\tilde{\boldsymbol{\mu}} &= -\partial_\mu \mathcal{F} = \partial \tilde{\boldsymbol{g}}^T \tilde{\boldsymbol{\Pi}}_y \tilde{\boldsymbol{\varepsilon}}_y + \partial_\mu \tilde{\boldsymbol{f}}^T \tilde{\boldsymbol{\Pi}}_x \tilde{\boldsymbol{\varepsilon}}_x - \mathcal{D}^T \tilde{\boldsymbol{\Pi}}_x \tilde{\boldsymbol{\varepsilon}}_x \\ \dot{\tilde{\boldsymbol{\nu}}} - \mathcal{D}\tilde{\boldsymbol{\nu}} &= -\partial_\nu \mathcal{F} = \partial_\nu \tilde{\boldsymbol{f}}^T \tilde{\boldsymbol{\Pi}}_x \tilde{\boldsymbol{\varepsilon}}_x - \tilde{\boldsymbol{\Pi}}_v \tilde{\boldsymbol{\varepsilon}}_v \end{aligned} \tag{32}$$

expressed in terms of precisions (or inverse variances) $\tilde{\boldsymbol{\Pi}}_y$, $\tilde{\boldsymbol{\Pi}}_x$, and $\tilde{\boldsymbol{\Pi}}_v$. Here, $\tilde{\boldsymbol{\varepsilon}}_s$, $\tilde{\boldsymbol{\varepsilon}}_x$, and $\tilde{\boldsymbol{\varepsilon}}_v$ respectively denote the prediction errors of sensory signals, hidden states dynamics, and prior over hidden causes:

$$\begin{aligned} \tilde{\boldsymbol{\varepsilon}}_y &= \tilde{\boldsymbol{y}} - \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{\mu}}) \\ \tilde{\boldsymbol{\varepsilon}}_x &= \mathcal{D}\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\nu}}) \\ \tilde{\boldsymbol{\varepsilon}}_v &= \tilde{\boldsymbol{\nu}} - \boldsymbol{\eta}_v \end{aligned} \tag{33}$$

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

$$\dot{\boldsymbol{a}} = -\partial_a \mathcal{F} = -\partial_a \tilde{\boldsymbol{g}}^T \tilde{\boldsymbol{\Pi}}_y \tilde{\boldsymbol{\varepsilon}}_y \tag{34}$$

This process selects the outcomes that correspond to the representation of the world currently maintained by the agent; in short, if during perceptual inference the agent's beliefs are updated so as to reflect the sensory signals, during *active inference* the agent acts with the goal of sampling those sensory signals that conform to its beliefs, as in a self-fulfilling prophecy. This sequence further minimizes the VFE, enabling the implementation of goal-directed behavior and maintaining the agent within predictable regimes of state space that are characteristic of the agent in question (e.g., avoiding collisions and reaching targets). [4].

4.2 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, its applicability is constrained as it struggles to accommodate broader classes of actions and decision-making processes. The VFE minimization can solely update the approximate posterior based on current or past observations, failing to incorporate possible future states and outcomes. To grant this ability to the agent, a quantity called the *Expected Free Energy* (EFE) is introduced [13, 14]. We define an active inference discrete model as a Partially Observable Markov decision processes (POMDPs):

$$p(\boldsymbol{s}_{1:T}, \boldsymbol{o}_{1:T}, \boldsymbol{\pi}) = p(\boldsymbol{s}_1) \cdot p(\boldsymbol{\pi}) \cdot \prod_{\tau=1}^{T} p(\boldsymbol{o}_\tau|\boldsymbol{s}_\tau) \cdot \prod_{\tau=2}^{T} p(\boldsymbol{s}_\tau|\boldsymbol{s}_{\tau-1}, \boldsymbol{\pi}) \tag{35}$$

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:

$$\mathcal{G}_\pi = \mathbb{E}\, q(\boldsymbol{s}, \boldsymbol{o}|\boldsymbol{\pi}) \left[\ln \frac{q(\boldsymbol{s}|\boldsymbol{\pi})}{p(\boldsymbol{s}, \boldsymbol{o}|\boldsymbol{\pi})} \right] \approx \underset{q(\boldsymbol{s},\boldsymbol{o}|\boldsymbol{\pi})}{\mathbb{E}} \left[\ln \frac{q(\boldsymbol{s})}{q(\boldsymbol{s}|\boldsymbol{o}, \boldsymbol{\pi})} \right] - \underset{q(\boldsymbol{o}|\boldsymbol{\pi})}{\mathbb{E}} [\ln p(\boldsymbol{o}|\boldsymbol{C})] \tag{36}$$

Here, $\boldsymbol{s}$ and $\boldsymbol{o}$ denote discrete states and outcomes, $\boldsymbol{\pi}$ the policies, τ a discrete time step. Notably, the policies are not mere stimulus-response mappings as in Reinforcement Learning schemes, but rather sequences of actions. Each of these elements can be expressed using categorical distributions:

$$\begin{aligned} p(\boldsymbol{s}_1) &= Cat(\boldsymbol{D}) & p(\boldsymbol{o}_\tau|\boldsymbol{s}_\tau) &= Cat(\boldsymbol{A}) \\ p(\boldsymbol{\pi}) &= Cat(\boldsymbol{E}) & p(\boldsymbol{s}_\tau|\boldsymbol{s}_{\tau-1}, \boldsymbol{\pi}) &= Cat(\boldsymbol{B}_{\pi,\tau}) \end{aligned} \tag{37}$$

where $\boldsymbol{D}$ 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. As before, exact computation of the posterior $p(\boldsymbol{s}_{1:T}, \boldsymbol{\pi}|\boldsymbol{o}_{1:T})$ is unfeasible given the intractability of the model evidence $p(\boldsymbol{o}_{1:T})$. Therefore, we resort to an approximate posterior distribution $q(\boldsymbol{s}_{1:T}, \boldsymbol{\pi})$, and minimize the Kullback-Leibler (KL) divergence between the approximate and real posteriors:

$$\begin{aligned} 0 \leq &= D_{KL}[q(\boldsymbol{s}_{1:T}, \boldsymbol{\pi})||p(\boldsymbol{s}_{1:T}, \boldsymbol{\pi}|\boldsymbol{o}_{1:T})] \\ &= \underset{q}{\mathbb{E}}[\ln q(\boldsymbol{s}_{1:T}, \boldsymbol{\pi}) - \ln p(\boldsymbol{o}_{1:T}, \boldsymbol{s}_{1:T}, \boldsymbol{\pi})] + \ln p(\boldsymbol{o}_{1:T}) \\ &= \mathcal{F} + \ln p(\boldsymbol{o}_{1:T}) \end{aligned} \tag{38}$$

Assuming a mean-field approximation, the approximate posterior factorizes into independent distributions:

$$\begin{aligned} p(\boldsymbol{s}_{1:T}|\boldsymbol{o}_{1:T}, \boldsymbol{\pi}) \approx q(\boldsymbol{s}_{1:T}, \boldsymbol{\pi}) &= q(\boldsymbol{\pi}) \prod_{\tau}^{T} q(\boldsymbol{s}_\tau|\boldsymbol{\pi}) \\ q(\boldsymbol{\pi}) &= Cat(\boldsymbol{\pi}) \\ q(\boldsymbol{s}_\tau|\boldsymbol{\pi}) &= Cat(\boldsymbol{s}_{\pi,\tau}) \end{aligned} \tag{39}$$

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 get a valid probability distribution:

$$\boldsymbol{s}_{\pi,\tau} = \sigma(\ln \boldsymbol{B}_{\pi,\tau-1}\boldsymbol{s}_{\pi,\tau-1} + \ln \boldsymbol{B}_{\pi,\tau}^T \boldsymbol{s}_{\pi,\tau+1} + \ln \boldsymbol{A}^T \boldsymbol{o}_\tau) \tag{40}$$

Likewise, to update $q(\boldsymbol{\pi})$, the approach involves combining messages from the policy's prior (expressed by matrix $\boldsymbol{E}$) and from future outcomes conditioned on policies – given that action planning requires the selection of policies that lead 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:

$$\mathcal{G}_\pi = \mathbb{E}\, q(\boldsymbol{s}, \boldsymbol{o}|\boldsymbol{\pi}) \left[\ln \frac{q(\boldsymbol{s}|\boldsymbol{\pi})}{p(\boldsymbol{s}, \boldsymbol{o}|\boldsymbol{\pi})} \right] \approx \underset{q(\boldsymbol{s},\boldsymbol{o}|\boldsymbol{\pi})}{\mathbb{E}} \left[\ln \frac{q(\boldsymbol{s})}{q(\boldsymbol{s}|\boldsymbol{o}, \boldsymbol{\pi})} \right] - \underset{q(\boldsymbol{o}|\boldsymbol{\pi})}{\mathbb{E}} [\ln p(\boldsymbol{o}|\boldsymbol{C})] \tag{41}$$

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. Under appropriate assumptions, the EFE conditioned at a specific time step τ becomes:

$$\begin{aligned} \mathcal{G}_\pi &\approx \sum_\tau D_{KL}[q(\boldsymbol{o}_\tau|\boldsymbol{\pi})||p(\boldsymbol{o}_\tau|\boldsymbol{C})] + \underset{q(\boldsymbol{s}_\tau|\boldsymbol{s}_{\tau-1},\boldsymbol{\pi})}{\mathbb{E}} [H[p(\boldsymbol{o}_\tau|\boldsymbol{s}_\tau)]] \\ &= \sum_\tau \boldsymbol{o}_{\pi,\tau}(\ln \boldsymbol{o}_{\pi,\tau} - \boldsymbol{C}_\tau) + \boldsymbol{s}_{\pi,\tau}\boldsymbol{H} \end{aligned} \tag{42}$$

where:

$$\boldsymbol{o}_{\pi,\tau} = \boldsymbol{A}\boldsymbol{s}_{\pi,\tau} \qquad \boldsymbol{C}_\tau = \ln p(\boldsymbol{o}_\tau|\boldsymbol{C}) \qquad \boldsymbol{H} = -diag(\boldsymbol{A}^T \ln \boldsymbol{A}) \tag{43}$$

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

$$\begin{aligned} \boldsymbol{\pi} &= \sigma(\ln \boldsymbol{E} - \boldsymbol{\mathcal{G}}) \\ u_t &= \arg\max_u \boldsymbol{\pi} \cdot [U_{\pi,t} = u] \end{aligned} \tag{44}$$

whose result in the same *self-evidencing* mechanism observed in the continuous framework.

4.3 Bayesian model reduction

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

$$p(\boldsymbol{\theta}, \boldsymbol{y}) = p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{45}$$

Consider then an additional distribution $p(\boldsymbol{y}, \boldsymbol{\theta}|m)$: this 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:

$$p(\boldsymbol{\theta}|\boldsymbol{y}, m) = p(\boldsymbol{\theta}|\boldsymbol{y})\frac{p(\boldsymbol{\theta}|m)p(\boldsymbol{y})}{p(\boldsymbol{\theta})p(\boldsymbol{y}|m)} \tag{46}$$

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:

$$p(\boldsymbol{y}|m) = p(\boldsymbol{y}) \int p(\boldsymbol{\theta}|\boldsymbol{y})\frac{p(\boldsymbol{\theta}|m)}{p(\boldsymbol{\theta})}d\boldsymbol{\theta} \tag{47}$$

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

$$\mathcal{F}[p(\boldsymbol{\theta}|m)] \approx \mathcal{F}[p(\boldsymbol{\theta})] + \ln \mathop{\mathbb{E}}_q \left[\frac{p(\boldsymbol{\theta}|m)}{p(\boldsymbol{\theta})}\right] \tag{48}$$

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:

$$\ln q(\boldsymbol{\theta}|m) = \ln q(\boldsymbol{\theta}) + \ln \frac{p(\boldsymbol{\theta}|m)}{p(\boldsymbol{\theta})} - \ln \mathop{\mathbb{E}}_q \left[\frac{p(\boldsymbol{\theta}|m)}{p(\boldsymbol{\theta})}\right] \tag{49}$$

The Laplace approximation [43] leads to a simple form of the approximate posterior and the reduced free energy. Assuming the following Gaussian distributions:

$$\begin{aligned} p(\boldsymbol{\theta}) &= \mathcal{N}(\boldsymbol{\eta}, \boldsymbol{\Sigma}) & q(\boldsymbol{\theta}) &= \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C}) \\ p(\boldsymbol{\theta}|m) &= \mathcal{N}(\boldsymbol{\eta}_m, \boldsymbol{\Sigma}_m) & q(\boldsymbol{\theta}|m) &= \mathcal{N}(\boldsymbol{\mu}_m, \boldsymbol{C}_m) \end{aligned} \tag{50}$$

The reduced free energy turns to:

$$\mathcal{F}[p(\boldsymbol{\theta}|m)] \approx \frac{1}{2}\ln|\boldsymbol{\Pi}_m \boldsymbol{P}\boldsymbol{C}_m\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{\mu}^T\boldsymbol{P}\boldsymbol{\mu} - \boldsymbol{\mu}_m^T\boldsymbol{P}_m\boldsymbol{\mu}_m - \boldsymbol{\eta}^T\boldsymbol{\Pi}\boldsymbol{\eta} + \boldsymbol{\eta}_m^T\boldsymbol{\Pi}_m\boldsymbol{\eta}_m) \tag{51}$$

expressed in terms of precisions of priors $\boldsymbol{P}$ and posteriors $\boldsymbol{\Pi}$. The reduced posterior mean and precision are computed through Equation 49:

$$\begin{aligned} \boldsymbol{\mu}_m &= \boldsymbol{C}_m(\boldsymbol{P}\boldsymbol{\mu} - \boldsymbol{\Pi}\boldsymbol{\eta} + \boldsymbol{\Pi}_m\boldsymbol{\eta}_m) \\ \boldsymbol{P}_m &= \boldsymbol{P} - \boldsymbol{\Pi} + \boldsymbol{\Pi}_m \end{aligned} \tag{52}$$

For a more detailed treatment on Bayesian model reduction, see [16, 17].

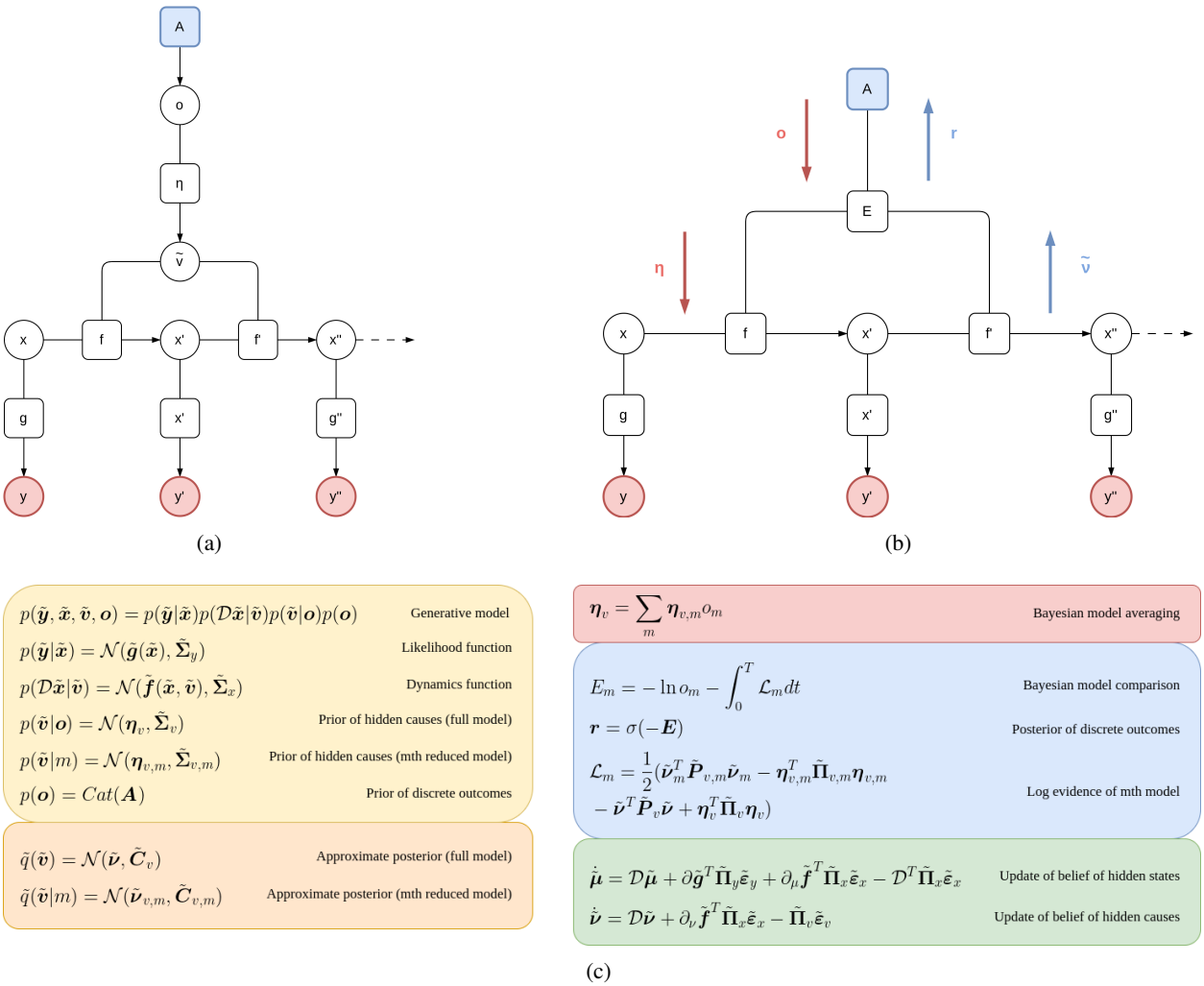


Figure 6: (a) Factor graph of a conventional hybrid model. Discrete outcomes $\boldsymbol{o}$ are generated by a prior $\boldsymbol{A}$, which we can assume coming from a discrete model with hidden states $\boldsymbol{s}$ – as in Section 4.2. In turn, the discrete outcomes act as a prior over the hidden causes $\tilde{\boldsymbol{v}}$ of the continuous model. (b) A graph highlighting the top-down (red) and bottom-up (blue) messages of Bayesian model reduction. The prior $\boldsymbol{\eta}$ is computed by Bayesian model averaging, which combines continuous priors $\boldsymbol{\eta}_{v,m}$ of specific outcome models (not displayed here) with the respective probabilities o_m. Bottom-up messages $\boldsymbol{r}$ entail a Bayesian model comparison, which compares the prior expectation of each model encoded in the surprise $-\ln o_m$ with the respective log evidence $\mathcal{L}_m$ accumulated over time. (c) Generative model, approximate posteriors, Bayesian model averaging, Bayesian model comparison, and update rule of the beliefs.

4.4 Hybrid models in active inference

Figure 6a shows the factor graph of a conventional hybrid model, whose generative model is assumed to be the following:

$$
\begin{aligned}
p(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}, \boldsymbol{o}) &= p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}})p(\mathcal{D}\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}})p(\tilde{\boldsymbol{v}}|\boldsymbol{o})p(\boldsymbol{o}) \\
p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{\Sigma}}_y) \\
p(\mathcal{D}\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}}) &= \mathcal{N}(\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}), \tilde{\boldsymbol{\Sigma}}_x) \\
p(\tilde{\boldsymbol{v}}|\boldsymbol{o}) &= \mathcal{N}(\boldsymbol{\eta}_v, \tilde{\boldsymbol{\Sigma}}_v) \\
p(\boldsymbol{o}) &= Cat(\boldsymbol{A})
\end{aligned}
\tag{53}
$$

where $\tilde{\boldsymbol{y}}$ are the continuous observations, $\tilde{\boldsymbol{g}}$ is the likelihood function, $\tilde{\boldsymbol{x}}$ are the continuous hidden states, $\tilde{\boldsymbol{f}}$ is the dynamics function, $\tilde{\boldsymbol{v}}$ are the hidden causes, $\boldsymbol{o}$ are the discrete outcomes with prior $\boldsymbol{A}$. For simplicity, we do not consider a complete discrete model as in Section 4.2, but assume that discrete outcomes can be generated by discrete hidden states at time τ through the likelihood matrix, i.e., $\boldsymbol{o}_\tau = \boldsymbol{A}\boldsymbol{s}_\tau$. This generative model is supposed to learn a representation of the environment modeled as in Section 4.1.

The probability distribution $p(\tilde{\boldsymbol{v}}|\boldsymbol{o})$ tells us that the hidden causes are generated by discrete outcomes $\boldsymbol{o}$ through a Gaussian distribution with mean $\boldsymbol{\eta}_v$; this is considered a full model over a complex set of hypotheses. We can infer the most likely hidden cause under the current discrete outcomes by assuming that the agent maintains M reduced models:

$$p(\tilde{\boldsymbol{v}}|m) = \mathcal{N}(\boldsymbol{\eta}_{v,m}, \tilde{\boldsymbol{\Sigma}}_{v,m}) \tag{54}$$

with the same likelihood of the full model but differing in their prior. This is a Gaussian mixture model, where the mixtures pertain to the generalized causes of a trajectory. In other words, we treat each outcome o_m as a particular model of dynamics, with prior $\boldsymbol{\eta}_{v,m}$ over the hidden causes. As highlighted in Figure 6b, the exchange of top-down and bottom-up messages between the discrete and continuous representations follows the derivations of Bayesian model reduction in Section 4.3. After having defined appropriate approximate posteriors, the continuous expectations in Equation 54 are computed through a Bayesian model average between the continuous priors of each outcome model, weighted by the respective prior probabilities:

$$\boldsymbol{\eta}_v = \sum_m \boldsymbol{\eta}_{v,m} o_m \tag{55}$$

Similarly, the probability distribution $\boldsymbol{r}$, over models, is found by first computing a Bayesian model comparison between the prior surprise and the log evidence accumulated over a continuous time T:

$$E_m = -\ln o_m - \int_0^T \mathcal{L}_m dt \tag{56}$$

where E_m and $\mathcal{L}_m$ are respectively the free energy and log evidence of the mth outcome model; then passing the free energy through a softmax function σ to get a proper probability, i.e., $\boldsymbol{r} = \sigma(-\boldsymbol{E})$. The log evidence approximates the reduced free energy of Equation 51, where the posterior is over hidden causes $\tilde{\boldsymbol{v}}$:

$$\begin{aligned} \mathcal{L}_m &= \ln p(\tilde{\boldsymbol{y}}|\boldsymbol{\eta}_{v,m}) - \ln p(\tilde{\boldsymbol{y}}|\boldsymbol{\eta}_v) \\ &= \frac{1}{2}(\tilde{\boldsymbol{\nu}}_m^T \tilde{\boldsymbol{P}}_{v,m} \tilde{\boldsymbol{\nu}}_m - \boldsymbol{\eta}_{v,m}^T \tilde{\boldsymbol{\Pi}}_{v,m} \boldsymbol{\eta}_{v,m} - \tilde{\boldsymbol{\nu}}^T \tilde{\boldsymbol{P}}_v \tilde{\boldsymbol{\nu}} + \boldsymbol{\eta}_v^T \tilde{\boldsymbol{\Pi}}_v \boldsymbol{\eta}_v) \end{aligned} \tag{57}$$

As a practical example, the reduced priors $\boldsymbol{\eta}_{v,m}$ could encode hypothetical positions of the agent's hand. Such reduced priors could be associated with different discrete hidden states and outcomes, encoding the probability that the defined positions will be observed (or realized). Via Equation 55, a discrete model can transform a high-level plan into a composite trajectory, e.g., for multi-step reaching [27]. Conversely, Equation 56 evaluates the probability of each reduced prior compared to the current posterior about the hand position. This message signals whether the agent has reached one of the positions, enabling the discrete model to compute the next reaching movement. See [25, 28] for more details on hybrid models.

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.

Code availability

Code and data are deposited in GitHub: https://github.com/priorelli/dynamic-inference

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] 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.

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

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

[12] 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.

[13] 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.

[14] 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.

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

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

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

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

[19] 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.

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

[21] 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.

[22] 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.

[23] 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.

[24] 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.

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

[26] 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.

[27] Thomas Parr, Jakub Limanowski, Vishal Rawji, and Karl Friston. The computational neurology of movement under active inference. *Brain*, 144(6):1799–1818, 2021.

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

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

[30] 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.

[31] 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.

[32] Matteo Priorelli and Ivilin Peev Stoianov. Dynamic planning in hierarchical active inference. *Neural Networks*, page 107075, January 2025.

[33] Matteo Priorelli and Ivilin Peev Stoianov. Slow but flexible or fast but rigid? discrete and continuous processes compared. *Heliyon*, page e39129, October 2024.

[34] Matteo Priorelli, Giovanni Pezzulo, and Ivilin Peev Stoianov. Deep kinematic inference affords efficient and scalable control of bodily movements. *PNAS*, 120, 2023.

[35] Matteo Priorelli, Ivilin Peev Stoianov, and Giovanni Pezzulo. Embodied decisions as active inference. *bioRxiv*, 2025.

[36] Nathan F. Lepora and Giovanni Pezzulo. Embodied choice: How action influences perceptual decision making. *PLOS Computational Biology*, 11(4):e1004110, April 2015.

[37] Christopher L Buckley and Taro Toyoizumi. A theory of how active behavior stabilises neural activity: Neural gain modulation by closed-loop environmental feedback. *PLoS computational biology*, 14(1):e1005926, 2018.

[38] Matteo Priorelli and Ivilin Peev Stoianov. Deep hybrid models: infer and plan in the real world. *arXiv*, 2024.

[39] Andy Clark. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204, 2013.

[40] Jakob Hohwy. *The Predictive Mind*. Oxford University Press UK, 2013.

[41] 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.

[42] Poppy Collis, Ryan Singh, Paul F Kinghorn, and Christopher L Buckley. Learning in hybrid active inference models, 2024.

[43] 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.