Deep kinematic inference affords efficient and scalable control of bodily movements

Matteo Priorelli[1], Giovanni Pezzulo[2], Ivilin Peev Stoianov,[1*]

[1]National Research Council, Institute of Cognitive Sciences and Technologies, Padova, Italy

[2]National Research Council, Institute of Cognitive Sciences and Technologies Rome, Italy

*Correspondence to: ivilinpeev.stoianov@cnr.it

Performing goal-directed movements requires mapping goals from extrinsic (workspace-relative) to intrinsic (body-relative) coordinates and then to motor signals. Mainstream approaches based on Optimal Control realize the mappings by minimizing cost functions, which is computationally demanding. Active inference instead uses generative models to produce sensory predictions, which allows computing motor signals through a cheaper inversion of the predictions. However, devising generative models to control complex kinematic plants like the human body is challenging. We introduce a novel Active Inference architecture that affords a simple but effective mapping from extrinsic to intrinsic coordinates via inference and easily scales up to drive complex kinematic plants. Rich goals can be specified in both intrinsic and extrinsic coordinates using both attractive and repulsive forces. The novel deep kinematic inference model reproduces sophisticated bodily movements and paves

the way for computationally efficient and biologically plausible control of actuated systems.

How do the brains of living systems support goal-directed movements implementing purposeful behavior? A standard assumption is that performing goal-directed actions requires two mappings reflecting perceptual and plant control processes (*1*). First, it is necessary to map goals and desired movement trajectories specified in *extrinsic* coordinates (e.g., in Cartesian space) into movements specified in *intrinsic* coordinates (e.g., as joint angles), via a process called *inverse kinematics*. Second, it is necessary to map joint angles into the torques and forces needed to move all the muscles of the body, via a process called *inverse dynamics*.

These two inversions (kinematic and dynamic) are computationally challenging. While computing the extrinsic representation of a given joint configuration is easy as the geometric mapping – the so-called *direct kinematics* – is univocal, finding the joint angles that correspond to a particular position is not straightforward. Computing a motor plan that brings the agent to that posture is even more difficult, since there are multiple possibilities of how the movement may be realized – and the computational complexity increases with the Degrees of Freedom (DoF) of the plant (e.g., a single arm versus the entire body).

Mainstream theories based on Optimal Control provide solutions to the two inversions that rest upon the optimization of a cost function (*2–6*). However, realizing certain movements such as handwriting or walking is difficult, since not all Bayes-optimal trajectories have well-defined cost functions (*7*). Furthermore, the required computations are usually challenging, since the model inversion requires considering the effects of the executed actions. These action effects are however perceived with delays that vary across and within sensory modalities (*8, 9*), which needs to be compensated using a forward model, which could introduce errors (*10*).

Active Inference offers an alternative scheme that does not require cost functions (*7, 11–15*). It proposes that agents are endowed with a generative model specifying the dynamics of their hidden states (e.g., hand positions over time) and that desired goals are encoded as priors over the dynamics (e.g., desired hand positions), which act as attracting states. Goal-directed movements are realized by first generating (proprioceptive) predictions from hidden states and then minimizing (proprioceptive) prediction errors, or the discrepancy between predicted and current sensations. Crucially, the mapping between proprioceptive predictions and control signals for the muscles is quite simple and can be implemented with minimal latency using reflex arcs in the spinal cord rather than requiring complex *inverse dynamics* computations (*13*). In fact, the inverse model now maps from proprioceptive sensations to actions, not from hidden states (in either intrinsic or extrinsic coordinates) to actions, as in Optimal Control (*10*). Prior expectations allow immediate first-guess response that is eventually refined by subsequent feedback (*1*).

While there is increasing interest in using Active Inference methods for biological control and robotics (*16–25*), their application in realistic settings has been limited so far, given two fundamental limitations. First, realizing generative models that afford effective kinematic inversions is challenging. State-of-the-art robotic implementations skip the kinematic inversion because they realize movements by only relying on intrinsic coordinates (*26, 27*). Despite its effectiveness, this scheme does not allow specifying goals in extrinsic coordinates and hence has several practical limitations. Other approaches embed classical kinematic inversion methods borrowed from Optimal Control, such as the Jacobian transpose (*10, 28*) or the pseudoinverse (*29*) directly within the dynamics of the hidden states. This has the disadvantage that extrinsic and intrinsic coordinates are mixed up and some computations are duplicated: extrinsic priors are mapped into intrinsic coordinates, which in turn generate extrinsic predictions to be compared with sensory observations – hence the extrinsic generative model has to be embedded also in the dynamics function.

Second, and crucially, the above systems do not model (i.e., maintain probabilistic beliefs about) the entire kinematic chain of the body to be controlled, but only the end effector (e.g., the final position of a hand). Therefore, they fall short of addressing naturalistic bodily movements that require the simultaneous coordination of multiple limbs and joints within a complex kinematic chain, such as that of the body of living organisms, and they also have limitations when addressing body movement restrictions due to obstacles.

Here, we introduce a novel Active Inference architecture that resolves the two above limitations and achieves effective and scalable motor control. First, we show that a generative model that maps intrinsic coordinates into extrinsic beliefs (IE model) affords an effective kinematic inversion through inference. Second, we show that replicating the scheme of the IE model to mimic the hierarchical structure of the agent's motor plant affords scalable control of bodily movements.

Results

Kinematic inversion through an Intrinsic-Extrinsic model

Active Inference can afford effective motor control through a kinematic generative model that decomposes latent states into two layers, which deal with intrinsic ($\boldsymbol{\mu}_i$) and extrinsic ($\boldsymbol{\mu}_e$) coordinates, respectively (see Fig. 1A). This "Intrinsic-Extrinsic (IE) model" has two appealing features compared to Optimal Control models and previous Active Inference implementations. First, the kinematic inversion – or the mapping from extrinsic to intrinsic coordinates – does not need to be specified in the model dynamics (*10, 28, 29*), but can be realized via inference, that is, by inverting the generative mapping from joint angles to Cartesian positions $\boldsymbol{\mu}_e = \boldsymbol{g}_e(\boldsymbol{\mu}_i)$.

Second, this scheme allows specifying priors at both intrinsic and extrinsic levels simultaneously. In robotics, it is common to specify task goals as attractors at the *extrinsic level* (e.g., a point to be reached). In this case, the attractor drives the inference of the most likely *intrinsic*

hidden state at the level above, which in turn affects movement execution through the pathway that goes from intrinsic hidden states to proprioception and finally to motor commands. Furthermore, specifying attractors at the *intrinsic* level is useful to deal with particular actions for which this is more natural or when extrinsic goals are difficult to define. For example, grasping actions can be realized by specifying priors over object-specific joint configurations (e.g., precision grip for small objects or power grip for large objects).

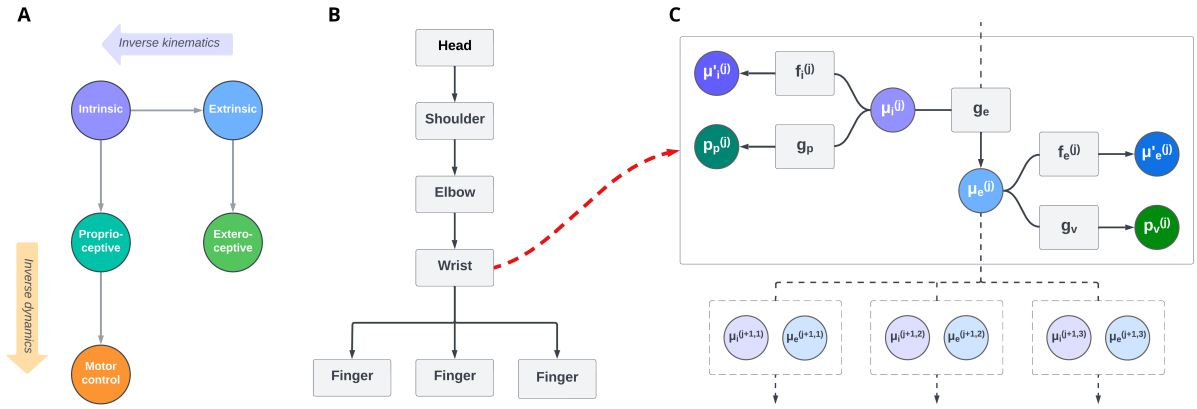


Figure 1: **Generative models for deep kinematic inference.** (**A**) The Intrinsic-Extrinsic (IE) generative model. The two hidden states are intrinsic and extrinsic coordinates, each with its own dynamics. The proprioceptive generative model extracts joint angles from the intrinsic belief, which are used for both posture inference and movement. At the same time, the exteroceptive model produces a visual prediction (in the simulations reported below, it is approximated by directly using the absolute position of the hand). Inverse kinematics is performed by inference, via backpropagation of extrinsic prediction errors. On the other hand, inverse dynamics from proprioceptive predictions to motor control signals is realized through reflex arcs. (**B**) The deep hierarchical generative model that extends the IE model. The deep model is exemplified for the control of a kinematic plant composed of seven blocks arranged hierarchically, from head to fingers (note that the bottom level is branched, as there are three fingers). Here, the highest (head) level only encodes an extrinsic offset, namely, the initial position and orientation of the origin's reference frame. (**C**) Factor graph of a single level of the deep hierarchical model. Each block has the same structure of the kinematic generative model shown in Panel A; however, intrinsic and extrinsic beliefs encode information about a single kinematic level, including absolute position and orientation, and generate proprioceptive and exteroceptive predictions only of that level. Furthermore, the extrinsic belief acts as a prior for all the layers below.

Extending the IE model to recapitulate the kinematic chain

The IE model introduced above affords efficient control but still performs direct kinematics in a single step; in other words, it only predicts the Cartesian coordinates of the end effector (e.g, the hand), not of the entire kinematic chain (e.g., the body). However, the IE model can be extended hierarchically in a straightforward manner, by noting that the direct kinematics of a complex motor plant can be always decomposed into a sequence of identical transformations (rotations and translations), one for each component of the plant. We exploit this fact to design a hierarchical generative model, composed of as many structurally identical blocks as the number of elements of the kinematic chain – where, crucially, each block is an IE model (see Fig. 1B-C). The hierarchical model affords direct kinematics by iterating the same operation across all the levels of the kinematic hierarchy, with a biologically plausible *local* message passing. Specifically, the Cartesian position of the tip of the body segment at level n is generated from the intrinsic belief over the body segment at the same level and the extrinsic belief over the body segment at the level above $n - 1$. Conversely, the (inverse kinematics) inference of the extrinsic belief at level n requires messages from the extrinsic belief at the level below $n + 1$. This biologically plausible form of local message passing is the same as hierarchical predictive coding for perceptual inference (*30,31*) – with top-down and bottom-up messages across blocks encoding *kinematic predictions* and *kinematic prediction errors*, respectively. Finally, it is worth noting that Fig. 1B shows a ramification at the bottom level: multiple fingers are attached to the same joint (the wrist). The ramifications can be treated in a straightforward manner, by assuming that the inference of the hidden states of the block at level n uses the average of the prediction errors generated by all the blocks at the level below $n + 1$ (*32*). See Methods for more information.

Defining goals for movement using attractive or repulsive forces

In Active Inference, goals for movement can be either imposed as priors or embedded into the dynamics function. The latter is usually composed of attractors that linearly minimize the distance between the current and desired states. To generate a dynamic goal, the desired state can depend on some components of the belief itself through a function $\boldsymbol{i}(\boldsymbol{\mu})$ embedding movement intentions (*33, 34*).

The deep hierarchical kinematic model introduced here permits formulating various kinds of goals at intrinsic and extrinsic levels; for example, the desired angle θ_d of a joint, the desired absolute position (x_d, y_d), or orientation ϕ_d of a limb. In addition, it is possible to define avoidance goals that correspond to repulsive forces, as shown in Methods. As for attractive forces, it is possible to define repulsive forces for each block of the kinematic plant, which ensures that every segment (rather than just the end effector) is pushed away from an obstacle. Finally, it is possible to define simultaneously multiple goals, such as reaching a particular position with the elbow while maintaining a vertical hand orientation, by combining multiple attractive and/or repulsive forces. This makes the deep model scalable and versatile, as we show in the next Section.

Applications of deep kinematic inference

Reaching task

We demonstrate the capacity of the deep hierarchical kinematic model in various motor tasks of increasing complexity. We start from the simple reaching task illustrated in Fig. 2A, which consists of moving a 4R robotic arm (blue) with realistic joint limits, to reach a static object (red). Here, we compare the performance of the deep hierarchical kinematic model (deep) with the simpler IE model (IE) and two alternative controllers based on Jacobian transpose (transp) and pseudoinverse (pinv).

The simulation results illustrated in Fig. 2B shows the performance of the four models for 1000 trials, with 500 steps for trial. For each trial, the beliefs of the models are initialized with random joint angles and segment positions (a more challenging scenario than starting from congruent intrinsic-extrinsic beliefs). Furthermore, a random target location is sampled and set as the reaching goal for the end effector. We evaluate the performance of the four models with regard to both perceptual inference (blue bars) and reaching ability (red bars), using various metrics (see Fig. 2B and Methods).

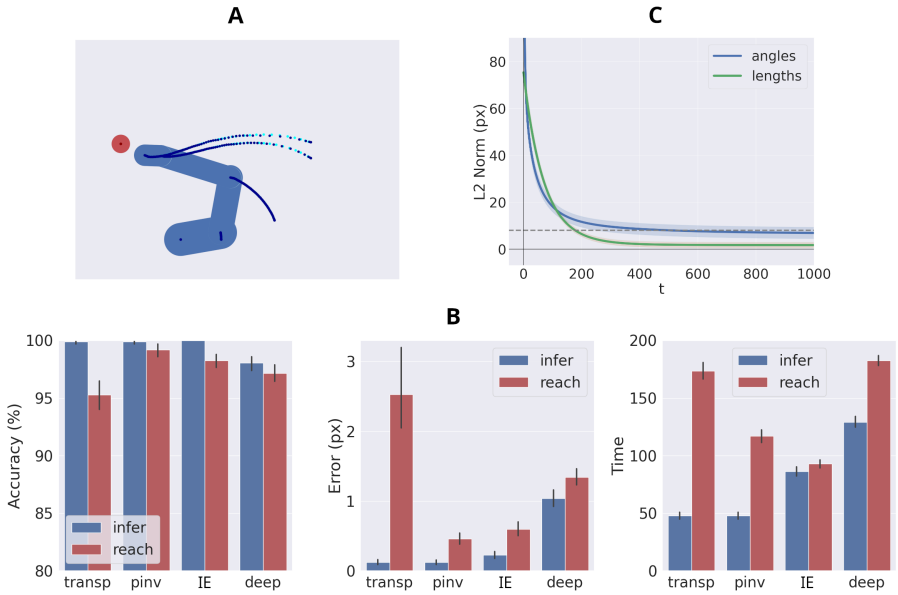


Figure 2: **Results.** (**A**) Simulation setup: a 4R robotic arm with realistic joint limits reaching a static target (red circle). The trajectory of each joint is shown with blue lines. (**B**) Comparison between the deep hierarchical kinematic model (deep), the simpler IE model (IE), and two standard Active Inference controllers based on Jacobian transpose (transp) and pseudoinverse (pinv). The blue and red bars denote the performance of the models during perceptual inference and reaching, respectively. Note that to better probe the model's ability to perform perceptual inference, we only allow the model to use visual observations, but not proprioceptive observations during this phase (however, the model can still use proprioceptive observations for reaching). (**C**) Evolution over time of the difference between true and estimated end effector positions (blue line), and between true and estimated lengths of the body segments (red line), aggregated over 1000 trials during inference only. The dashed green line represents the minimum distance defining a successful inference.

The performance of the four models is near-optimal during both perceptual inference and reaching. The rare failures are trials in which the models find impossible trajectories and therefore cannot reach the target configurations (note that we did not include any prior over the direction of movement). The performance of the deep model is on par with the other models. The slightly lower performance compared to the IE model is due to the fact that the deep model requires more time to propagate messages across levels, resulting in failed trials when the time steps are fixed. The Jacobian pseudoinverse method performs slightly better than the other models, but the transpose method has a much higher final error. Nevertheless, in all cases the average final error is below the minimum distance to consider a trial successful (set to 8 pixels).

Learning and adaptation of the motor plant

Interestingly, unlike the other models, the deep model includes beliefs about the length of its body segments and hence can infer them over time, based on sensory observations. This affords the rapid adaptation of the agent to changes in the motor plant, e.g., when using a tool that increases the length of the last segment or adds a new joint at a specific position. To assess this capability, we performed an experiment in which the beliefs about both angles and body segments are randomly initialized. We evaluated the task over 1000 trials, by comparing the estimated positions of every segment (computed from the joint angles belief) and the estimated lengths with the real values of the same variables. The simulation results illustrated in 2C show that the deep model successfully infers the length of its body segments (green line) in addition to inferring its joint angles (blue line), within a single trial.

In sum, these results show that the deep model does not lose efficiency compared to previous implementations and it additionally permits self-modeling. The next simulations will show that different from previous methods, it can be scaled up easily to complex motor plants.

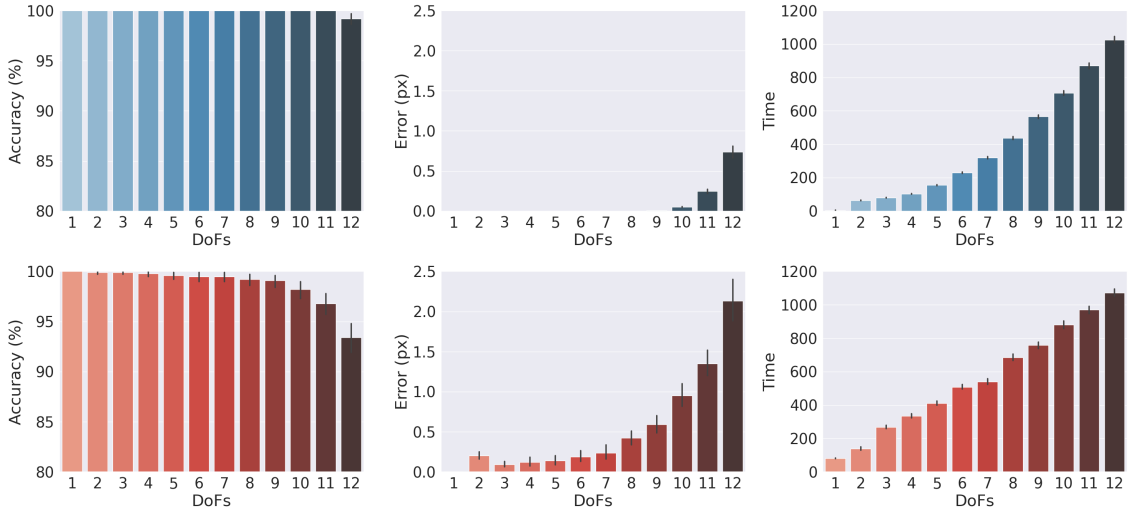


Figure 3: Comparison between hierarchical models with increasing DoF (here, plant depth). Inference (top) and reach (bottom) simulations.

Deep inference

Fig. 3 shows an evaluation of the deep hierarchical model on a robotic arm like that in Fig. 2 but with an increasing number of joints and equal segment lengths. This simulation was assessed on 1000 trials, with 2000 time steps for every trial. Since we removed the angles limits, the performances are optimal for inference and near-optimal for action (some trials fail, because the time needed to backpropagate the attractor toward the deepest levels exceeds the time limits, which affects both accuracy and mean error).

More complex control tasks

In addition to the previous simulations, the deep model can also control simpler (e.g., arm-like) and more complex (body-like) plants in a variety of tasks, which require simultaneously tracking a dynamic object (red) while avoiding another dynamic object (green) (Fig. 4A), making lateral movements while maintaining a vertical orientation of the last segment (Fig. 4B), performing circular movements in Cartesian space, where attractors are trajectories defined in

extrinsic coordinates (Fig. 4C), avoiding a dynamic object with a human-like body (Fig. 4D), reaching simultaneously multiple targets with different limbs of a human-like body (Fig. 4E), or of a tree-like body with 28 DoF and multiple ramifications (Fig. 4F); see the Supplementary Materials for more information about these and various other control tasks.

Crucially, the same deep model can achieve all these (and other) tasks by simply specifying different goals, using attractive or repulsive forces, or a combination of both, without having to define ad-hoc cost functions. Compared to other approaches, the deep model is therefore particularly easy to scale up to address complex kinematic structures and control problems with multiple goals.

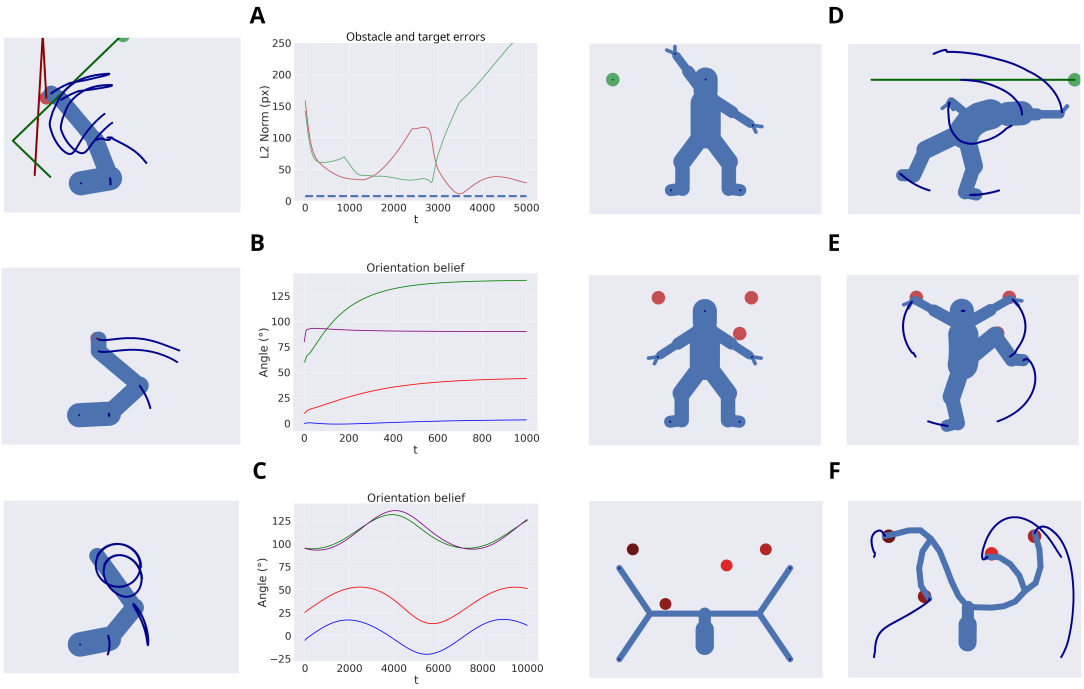


Figure 4: **Applications.** Track and avoid (**A**), maintain orientation (**B**), perform circular trajectory (**C**), avoid with a full human body (**D**), reach multiple targets with a full human body (**E**), and with a more complex kinematic tree (**F**). More information about the specific tasks and other applications can be found in the Supplementary Materials.

Discussion

Active Inference proposes that goal-directed movements are realized by using a generative model of the coupled agent-environment system to produce predictions and then minimizing prediction errors resulting from discrepancies between predicted and current sensations. Despite its strong theoretical appeal (*1*), this formulation has enjoyed limited applications in motor control and robotics, given the challenge of specifying appropriate generative models to map different (extrinsic and intrinsic) coordinates of movement and to control complex kinematic structures, such as the human body.

Here, we show that a generative (IE) model that maintains distinct beliefs over intrinsic and extrinsic coordinates affords effective control of goal-directed movements, because it allows defining intrinsic and extrinsic goals as attractor states and solving inverse kinematics via (active) inference, which is typically challenging in other frames like Optimal Control. Furthermore, we show that a deep hierarchical extension of the IE scheme drastically improves its scalability. In particular, replicating the same IE generative model for each block of the kinematic chain simplifies the computations of the direct kinematics and permits controlling complex kinematic structures that require the simultaneous coordination of multiple segments.

The proposed deep hierarchical architecture has four main advantages compared to Optimal Control schemes or state-of-the-art models in the Active Inference literature. First, it is efficient: performing the kinematic inversion through inference drastically reduces its computational cost. Second, it is scalable: it is possible to design generative models for sophisticated kinematic chains by simply connecting multiple IE models. Since the overall generative model always mimics the entire kinematic chain, adding or removing novel segments to it (e.g., to model tool use) simply requires adding or removing the corresponding blocks to the overall generative model. Third, it affords self-modeling (*35*): the model continuously infers the structure of the

motor plant online and can rapidly re-adapt to changes in (for example) limb lengths. Fourth, it affords a biologically plausible form of inference that only uses *local* message passing and *asynchronous* computations, akin to hierarchical predictive coding for perception (*11,30,31,36*).

The functioning of the deep kinematic model, which replicates the same computational block across various levels, has some analogies with deep neural networks. It is nowadays common to use deep networks (e.g., variational autoencoders) as generative models for Active Inference agents (*26, 33*). Despite their effectiveness, deep networks are treated as black boxes during the free energy minimization: the belief over hidden states only receives a single gradient and it is unaware of the internal computations performed by the backpropagation algorithm. This has the downside that the weights of the generative model have to be learned a priori and remain constant throughout the Active Inference process, hence the agent cannot adapt them when it receives novel sensory observations. Furthermore, the agent has no control over the dynamics of the generative model and thus it cannot control the intermediate levels of the deep network toward potentially preferred states. In other words, almost all the hard work is delegated to the deep network and the only job left to the Active Inference agent is to infer the highest level of the hidden states. In contrast, one appeal of Predictive Coding methods is that they can simultaneously perform hidden state inference and learning (*11, 30, 31*). Designing a hierarchical structure that uses the same rules of free energy (or prediction error) minimization throughout all the levels means that the agent can constantly modify its internal models to match both sensory observations and prior expectations, maintaining accurate and up-to-date predictions everywhere in the control system.

Although we exemplified a deep kinematic inference for the control of human body movements, the same design could be used to control other, potentially more sophisticated plants, composed of multiple ramifications. The deep kinematic inference approach can therefore pave the way for the efficient and biologically plausible control of general actuated systems.

Methods

Hierarchical Active Inference

The core of Active Inference is the use of a generative model to generate predictions and minimize any prediction error resulting from discrepancies between predictions and observations. The generative model depends on three elements encoded in generalized coordinates: hidden states $\tilde{\boldsymbol{x}}$, hidden causes $\tilde{\boldsymbol{v}}$ and sensory signals $\tilde{\boldsymbol{s}}$. These elements are related by a nonlinear system specifying the generation of sensory signals and the evolution of latent states over time:

$$\begin{aligned} \tilde{\boldsymbol{s}} &= \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}) + \boldsymbol{w}_s \\ \mathcal{D}\tilde{\boldsymbol{x}} &= \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}) + \boldsymbol{w}_x \end{aligned} \tag{1}$$

The associated joint probability is usually factored into independent distributions:

$$p(\tilde{\boldsymbol{s}}, \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}) = p(\tilde{\boldsymbol{s}}|\tilde{\boldsymbol{x}})p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}})p(\tilde{\boldsymbol{v}}) \tag{2}$$

where each distribution is usually approximated by Gaussian functions:

$$\begin{aligned} p(\tilde{\boldsymbol{s}}|\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{\Pi}}_s^{-1}) \\ p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}}) &= \mathcal{N}(\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}), \tilde{\boldsymbol{\Pi}}_x^{-1}) \\ p(\tilde{\boldsymbol{v}}|\eta) &= \mathcal{N}(\eta, \tilde{\boldsymbol{\Pi}}_v^{-1}) \end{aligned} \tag{3}$$

Following a variational inference approach (*37*), these distributions are inferred through approximate posteriors $q(\tilde{\boldsymbol{x}})$ and $q(\tilde{\boldsymbol{v}})$. Minimizing the Variational Free Energy (VFE) $\mathcal{F}$, defined as the difference between the KL divergence of the (real and approximate) posteriors and the log evidence:

$$\mathcal{F} = \underset{q(\tilde{\boldsymbol{x}})}{E}\left[\ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})}\right] = \underset{q(\tilde{\boldsymbol{x}})}{E}\left[\ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})}\right] - \ln p(\tilde{\boldsymbol{y}}) \tag{4}$$

reduces to the minimization of prediction errors, and the updates of the beliefs $\tilde{\boldsymbol{\mu}}$ and $\tilde{\boldsymbol{\nu}}$ over the latent states become:

$$\begin{aligned} \dot{\tilde{\boldsymbol{\mu}}} - \mathcal{D}\tilde{\boldsymbol{\mu}} &= -\partial_\mu \mathcal{F} = \partial \tilde{\boldsymbol{g}}^T \tilde{\boldsymbol{\Pi}}_s \tilde{\boldsymbol{\varepsilon}}_s + \partial_\mu \tilde{\boldsymbol{f}}^T \tilde{\boldsymbol{\Pi}}_\mu \tilde{\boldsymbol{\varepsilon}}_\mu - \mathcal{D}^T \tilde{\boldsymbol{\Pi}}_\mu \tilde{\boldsymbol{\varepsilon}}_\mu \\ \dot{\tilde{\boldsymbol{\nu}}} - \mathcal{D}\tilde{\boldsymbol{\nu}} &= -\partial_\nu \mathcal{F} = \partial_\nu \tilde{\boldsymbol{f}}^T \tilde{\boldsymbol{\Pi}}_\mu \tilde{\boldsymbol{\varepsilon}}_\mu - \tilde{\boldsymbol{\Pi}}_\nu \tilde{\boldsymbol{\varepsilon}}_v \end{aligned} \tag{5}$$

where $\tilde{\varepsilon}_s$, $\tilde{\varepsilon}_\mu$ and $\tilde{\varepsilon}_\nu$ are respectively the prediction errors of sensory signal, dynamics and prior:

$$\begin{aligned}
\tilde{\varepsilon}_s &= \tilde{\boldsymbol{s}} - \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{\mu}}) \\
\tilde{\varepsilon}_\mu &= \mathcal{D}\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\nu}}) \\
\tilde{\varepsilon}_\nu &= \tilde{\boldsymbol{\nu}} - \eta
\end{aligned} \tag{6}$$

A simple formulation of Active Inference can deal with several tasks, but its main strength relies on a hierarchical structure that allows the brain to learn the inherently hierarchical relations between sensory observations and their causes (*38*).

This structure can be easily scaled up by linking every hidden cause with another generative model; thus, the prior becomes the prediction from the layer above, while the observation becomes the likelihood of the layer below:

$$\begin{aligned}
\dot{\tilde{\boldsymbol{\mu}}}^{(i)} &= \mathcal{D}\tilde{\boldsymbol{\mu}}^{(i)} + \partial\tilde{\boldsymbol{g}}^{(i)T}\tilde{\boldsymbol{\Pi}}_\nu^{(i)}\tilde{\varepsilon}_\nu^{(i)} + \partial_\mu\tilde{\boldsymbol{f}}^{(i)T}\tilde{\boldsymbol{\Pi}}_\mu^{(i)}\tilde{\varepsilon}_\mu^{(i)} - \mathcal{D}^T\tilde{\boldsymbol{\Pi}}_\mu^{(i)}\tilde{\varepsilon}_\mu^{(i)} \\
\dot{\tilde{\boldsymbol{\nu}}}^{(i)} &= \mathcal{D}\tilde{\boldsymbol{\nu}}^{(i)} + \partial_\nu\tilde{\boldsymbol{f}}^{(i)T}\tilde{\boldsymbol{\Pi}}_\mu^{(i)}\tilde{\varepsilon}_\mu^{(i)} - \tilde{\boldsymbol{\Pi}}_\nu^{(i-1)}\tilde{\varepsilon}_\nu^{(i-1)}
\end{aligned} \tag{7}$$

where:

$$\begin{aligned}
\tilde{\varepsilon}_\mu^{(i)} &= \mathcal{D}\tilde{\boldsymbol{\mu}}_\mu^{(i)} - \tilde{\boldsymbol{f}}^{(i)}(\tilde{\boldsymbol{\mu}}^{(i)}, \tilde{\boldsymbol{\nu}}^{(i)}) \\
\tilde{\varepsilon}_\nu^{(i)} &= \tilde{\boldsymbol{\mu}}_\nu^{(i+1)} - \tilde{\boldsymbol{g}}^{(i)}(\tilde{\boldsymbol{\mu}}^{(i)})
\end{aligned} \tag{8}$$

On the other hand, action is realized by minimizing the proprioceptive component of the VFE with respect to the motor control signals $\boldsymbol{a}$:

$$\dot{\boldsymbol{a}} = -\partial_a\mathcal{F}_p = -\partial_a\tilde{\boldsymbol{s}}_p\tilde{\boldsymbol{\Pi}}_p\tilde{\varepsilon}_p \tag{9}$$

where $\partial_a\boldsymbol{s}_p$ is the partial derivative of proprioceptive observations with respect to the motor control signals, $\tilde{\boldsymbol{\Pi}}_p$ are the precisions of proprioceptive generative models, and $\tilde{\varepsilon}_p$ are the generalized proprioceptive prediction errors:

$$\tilde{\varepsilon}_p = \tilde{\boldsymbol{s}}_p - \tilde{\boldsymbol{g}}_p(\tilde{\boldsymbol{\mu}}) \tag{10}$$

Kinematic models: belief update

In the IE model (shown in Fig. S1), the intrinsic belief encodes every joint angle in parallel while the extrinsic belief only encodes the Cartesian position of the end effector. For a 2R robotic arm, such beliefs are related by:

$$\begin{aligned} \boldsymbol{\mu}_i &= [\theta_1, \theta_2] \\ \boldsymbol{\mu}_e &= [x_e, y_e] = \boldsymbol{g}_e(\boldsymbol{\mu}_i) \\ &= [l_1 c_1 + l_2 c_{12}, l_1 s_1 + l_2 s_{12}] \end{aligned} \tag{11}$$

The update equation for the intrinsic belief is then a combination of proprioceptive likelihood, extrinsic likelihood, and intrinsic attractor dynamics:

$$\dot{\tilde{\boldsymbol{\mu}}}_i = \mathcal{D}\tilde{\boldsymbol{\mu}}_i - \partial_{\mu_i}\mathcal{F} = \begin{bmatrix} \boldsymbol{\mu}_i' + \partial \boldsymbol{g}_p^T \pi_p \boldsymbol{\varepsilon}_p + \partial \boldsymbol{g}_e^T \pi_e \boldsymbol{\varepsilon}_e + \partial \boldsymbol{f}_i^T \pi_{\mu_i} \boldsymbol{\varepsilon}_{\mu_i} \\ -\pi_{\mu_i} \boldsymbol{\varepsilon}_{\mu_i} \end{bmatrix} \tag{12}$$

The kinematic inversion is performed through the gradient $\partial \boldsymbol{g}_e$:

$$\partial \boldsymbol{g}_e = \begin{bmatrix} -l_1 s_1 - l_2 s_{12} \\ l_1 c_1 + l_2 c_{12} \end{bmatrix} \tag{13}$$

where we used a more compact notation to indicate the sine and cosine of the angles:

$$\begin{aligned} c_1 &= \cos(\theta_1) & c_{12} &= \cos(\theta_1 + \theta_2) = c_1 c_2 - s_1 s_2 \\ s_1 &= \sin(\theta_1) & s_{12} &= \sin(\theta_1 + \theta_2) = c_1 s_2 + s_1 c_2 \end{aligned} \tag{14}$$

Similarly, the update for the extrinsic belief is a combination of intrinsic prior, visual likelihood, and extrinsic attractor dynamics:

$$\dot{\tilde{\boldsymbol{\mu}}}_e = \mathcal{D}\tilde{\boldsymbol{\mu}}_e - \partial_{\mu_e}\mathcal{F} = \begin{bmatrix} \boldsymbol{\mu}_e' - \pi_e \boldsymbol{\varepsilon}_e + \partial \boldsymbol{g}_v^T \pi_v \boldsymbol{\varepsilon}_v + \partial \boldsymbol{f}_e^T \pi_{\mu_e} \boldsymbol{\varepsilon}_{\mu_e} \\ -\pi_{\mu_e} \boldsymbol{\varepsilon}_{\mu_e} \end{bmatrix} \tag{15}$$

Concerning the deep hierarchical model, if we consider a simple 2D arm, each block could include an intrinsic belief over pairs of joint angles and segment lengths $\boldsymbol{\mu}_i^{(j)} = (\theta^{(j)}, l^{(j)})$ and an extrinsic belief over the position of the segment's extremity and its absolute orientation

$\boldsymbol{\mu}_e^{(j)} = (x^{(j)}, y^{(j)}, \phi^{(j)})$. The resulting kinematic generative model is:

$$\boldsymbol{\mu}_e^{(j)} = \boldsymbol{g}_e(\boldsymbol{\mu}_i^{(j)}, \boldsymbol{\mu}_e^{(j-1)}) = \begin{bmatrix} x^{(j-1)} + l^{(j)} c_{\theta,\phi}^{(j)} \\ y^{(j-1)} + l^{(j)} s_{\theta,\phi}^{(j)} \\ \phi^{(j-1)} + \theta^{(j)} \end{bmatrix} \tag{16}$$

The update of the intrinsic belief is equivalent to the IE model:

$$\dot{\tilde{\boldsymbol{\mu}}}_i^{(j)} = \begin{bmatrix} \boldsymbol{\mu}_i'^{(j)} + \partial \boldsymbol{g}_p^T \pi_p^{(j)} \boldsymbol{\varepsilon}_p^{(j)} + \partial_{\mu_i} \boldsymbol{g}_e^T \pi_e^{(j+1)T} \boldsymbol{\varepsilon}_e^{(j+1)} + \partial \boldsymbol{f}_i^{(j)T} \pi_{\mu_i}^{(j)} \boldsymbol{\varepsilon}_{\mu_i}^{(j)} \\ -\pi_{\mu_i}^{(j)} \boldsymbol{\varepsilon}_{\mu_i}^{(j)} \end{bmatrix} \tag{17}$$

However, the gradient of the kinematic generative model (expressed by Equation 16) with respect to the intrinsic belief is simpler, since it only depends on the intrinsic components of that level:

$$\frac{\partial \boldsymbol{g}_e}{\partial \boldsymbol{\mu}_i^{(j)}} = \begin{bmatrix} \partial_\theta \boldsymbol{g}_e \\ \partial_l \boldsymbol{g}_e \end{bmatrix} = \begin{bmatrix} -l^{(j)} s_{\theta,\phi}^{(j)} & l^{(j)} c_{\theta,\phi}^{(j)} & 1 \\ c_{\theta,\phi}^{(j)} & s_{\theta,\phi}^{(j)} & 0 \end{bmatrix} \tag{18}$$

where the second row of the gradient (18) allows inferring and learning segment lengths.

The update of the extrinsic belief still includes every component of the IE model, but with the addition of the extrinsic likelihood from the next layer $\partial_e \boldsymbol{g}_e$, which in this case is the sum of all the segments it is attached to:

$$\dot{\tilde{\boldsymbol{\mu}}}_e^{(j)} = \begin{bmatrix} \boldsymbol{\mu}_e'^{(j)} - \pi_e^{(j)} \boldsymbol{\varepsilon}_e^{(j)} + \sum_m \partial_{\mu_e} \boldsymbol{g}_e^T \pi_e^{(j+1,m)} \boldsymbol{\varepsilon}_e^{(j+1,m)} + \partial \boldsymbol{g}_v^T \pi_v^{(i)} \boldsymbol{\varepsilon}_v^{(i)} + \partial \boldsymbol{f}_e^{(j)T} \pi_{\mu_e}^{(j)} \boldsymbol{\varepsilon}_{\mu_e}^{(j)} \\ -\pi_{\mu_e}^{(j)} \boldsymbol{\varepsilon}_{\mu_e}^{(j)} \end{bmatrix} \tag{19}$$

The gradient of this new term connects every layer in the hierarchy since intrinsic beliefs communicate with their extrinsic predictions by:

$$\frac{\partial \boldsymbol{g}_e}{\partial \boldsymbol{\mu}_e^{(j)}} = \begin{bmatrix} \partial_x \boldsymbol{g}_e \\ \partial_y \boldsymbol{g}_e \\ \partial_\phi \boldsymbol{g}_e \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -l^{(j)} s_{\theta,\phi}^{(j)} & l^{(j)} c_{\theta,\phi}^{(j)} & 1 \end{bmatrix} \tag{20}$$

Ramifications

If we consider a ramification with M extrinsic outputs $\boldsymbol{\mu}_e^{(j+1,m)} = \boldsymbol{g}_e(\boldsymbol{\mu}_i^{(j+1)}, \boldsymbol{\mu}_e^{(j)})$, the update equation for the parent node $\boldsymbol{\mu}_e^{(j)}$ is proportional to:

$$\dot{\boldsymbol{\mu}}_e^{(j)} \propto \sum_m^M \partial_{\mu_e} \boldsymbol{g}_e^T \pi_e^{(j+1,m)} \boldsymbol{\varepsilon}_e^{(j+1,m)} \tag{21}$$

where $\boldsymbol{\varepsilon}_e^{(j+1,m)}$ is the extrinsic prediction error of block $(j+1, m)$ and $\pi_e^{(j+1,m)}$ is the corresponding precision. The latter thus weighs the contribution for each branch of the level.

Defining attractive goals and repulsive forces

Goals can be formulated at intrinsic and extrinsic levels; they could specify, for example, the desired angle θ_d of a joint, the desired absolute position (x_d, y_d), or orientation ϕ_d of a limb:

$$\begin{aligned} \boldsymbol{h}_i^{(j)} &= \boldsymbol{i}_\theta(\boldsymbol{\mu}_i^{(j)}) = \begin{bmatrix}\theta_d & l^{(j)}\end{bmatrix} \\ \boldsymbol{h}_e^{(j)} &= \boldsymbol{i}_p(\boldsymbol{\mu}_e^{(j)}) = \begin{bmatrix}x_d & y_d & \phi^{(j)}\end{bmatrix} \\ \boldsymbol{h}_e^{(j)} &= \boldsymbol{i}_\phi(\boldsymbol{\mu}_e^{(j)}) = \begin{bmatrix}x^{(j)} & y^{(j)} & \phi_d\end{bmatrix} \end{aligned} \tag{22}$$

Attractors linearly minimize the distance between the current and desired states, i.e., $\boldsymbol{f}_a(\boldsymbol{\mu}) = k_a \cdot \boldsymbol{e}_a$, where k_a is an attractive gain and $\boldsymbol{e}_a = \boldsymbol{\mu} - \boldsymbol{\mu}_d$

In turn, following the Artificial Field Potentials theory, repulsive forces can be defined as:

$$\boldsymbol{f}_r(\boldsymbol{\mu}) = \begin{cases} k_r \cdot (1/\mu^* - 1/||\boldsymbol{e}_r||) \cdot \boldsymbol{e}_r/||\boldsymbol{e}_r||^3, & \text{if } ||\boldsymbol{e}_r|| \leq \mu^* \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

where k_r is a repulsive gain, μ^* is the range of influence of the repulsive force, and $\boldsymbol{e}_r = \boldsymbol{\mu} - \boldsymbol{\mu}_r$ is the difference between the belief and repulsive state.

Assessment Metrics

The three metrics used to evaluate perceptual inference are: (i) *perception accuracy*: success in finding a joint configuration corresponding to the true target location within 8 pixels; (ii) *perception error*: L^2 distance between the true and estimated target positions at the end of the trial; (iii) *perception time*: number of time steps needed to successfully estimate the target position.

The three metrics used to evaluate reaching are (iv) *reach accuracy*: success in approaching the target within 8 pixels; (v) *reach error*: L^2 end effector-target distance at the end of the

trial; (vi) *reach time*: number of time steps needed for the end effector to reach the target in successful trials. Note that the reaching task is more challenging than the inference task, since it requires inferring the final arm configuration based on information about only the position of the last segment.

References

1. Floegel, M., Kasper, J., Perrier, P. & Kell, C. A. How the conception of control influences our understanding of actions. *Nature Reviews Neuroscience* **24**, 313–329 (2023).

2. Todorov, E. Optimality principles in sensorimotor control. *Nature Neuroscience* **7**, 907–915 (2004).

3. Diedrichsen, J., Shadmehr, R. & Ivry, R. B. The coordination of movement: optimal feedback control and beyond. *Trends in cognitive sciences* **14**, 31–39 (2010).

4. Shadmehr, R., Smith, M. A. & Krakauer, J. W. Error correction, sensory prediction, and adaptation in motor control. *Annual review of neuroscience* **33**, 89–108 (2010).

5. McNamee, D. & Wolpert, D. M. Internal models in biological control. *Annual review of control, robotics, and autonomous systems* **2**, 339–364 (2019).

6. Stengel, R. F. *Optimal control and estimation* (Courier Corporation, 1994).

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

8. Franklin, D. W. & Wolpert, D. M. Computational mechanisms of sensorimotor control. *Neuron* **72**, 425–442 (2011). URL http://dx.doi.org/10.1016/j.neuron.2011.10.006.

9. Hayhoe, M. M. Vision and Action. *Annual Review of Vision Science* **3**, 389–413 (2017).

10. Friston, K. J., Daunizeau, J., Kilner, J. & Kiebel, S. J. Action and behavior: A free-energy formulation. *Biological Cybernetics* **102**, 227–260 (2010).

11. Parr, T., Pezzulo, G. & Friston, K. J. *Active inference: the free energy principle in mind, brain, and behavior* (MIT Press, 2022).

12. Pio-Lopez, L., Nizard, A., Friston, K. & Pezzulo, G. Active inference and robot control: a case study. *Journal of The Royal Society Interface* **13**, 20160616 (2016).

13. Adams, R. A., Shipp, S. & Friston, K. J. Predictions not commands: Active inference in the motor system. *Brain Structure and Function* **218**, 611–643 (2013).

14. Baltieri, M. & Buckley, C. L. *Active inference: Computational models of motor control without efference copy* (ACM, New York, 2019).

15. Lanillos, P. & Cheng, G. Adaptive Robot Body Learning and Estimation Through Predictive Coding. *IEEE International Conference on Intelligent Robots and Systems* 4083–4090 (2018).

16. Lanillos, P. *et al.* Active Inference in Robotics and Artificial Agents: Survey and Challenges. *arXiv* 1–20 (2021). URL http://arxiv.org/abs/2112.01871.

17. Ciria, A., Schillaci, G., Pezzulo, G., Hafner, V. V. & Lara, B. Predictive processing in cognitive robotics: a review. *Neural Computation* **33**, 1402–1432 (2021).

18. Maselli, A., Lanillos, P. & Pezzulo, G. Active inference unifies intentional and conflict-resolution imperatives of motor control. *PLoS computational biology* **18**, e1010095 (2022).

19. Da Costa, L., Lanillos, P., Sajid, N., Friston, K. & Khan, S. How active inference could help revolutionise robotics. *Entropy* **24**, 361 (2022).

20. Mannella, F., Maggiore, F., Baltieri, M. & Pezzulo, G. Active inference through whiskers. *Neural Networks* **144**, 428–437 (2021).

21. Ahmadi, A. & Tani, J. A Novel Predictive-Coding-Inspired Variational RNN Model for Online Prediction and Recognition. *Neural Computation* **31**, 2025–2074 (2019).

22. Pezzato, C., Ferrari, R. & Corbato, C. H. A Novel Adaptive Controller for Robot Manipulators Based on Active Inference. *IEEE Robotics and Automation Letters* **5**, 2973–2980 (2020).

23. Mazzaglia, P., Verbelen, T., Catal, O. & Dhoedt, B. The Free Energy Principle for Perception and Action : A Deep Learning Perspective. *Entropy* (2021).

24. Catal, O., Verbelen, T., Van de Maele, T., Dhoedt, B. & Safron, A. Robot navigation as hierarchical active inference. *Neural Networks* **142**, 192–204 (2021).

25. Matsumoto, T. & Tani, J. Goal-directed planning for habituated agents by active inference using a variational recurrent neural network. *Entropy* **22**, 564 (2020).

26. Sancaktar, C., van Gerven, M. A. J. & Lanillos, P. *End-to-End Pixel-Based Deep Active Inference for Body Perception and Action*, 1–8 (2020). 2001.05847.

27. Baioumy, M., Duckworth, P., Lacerda, B. & Hawes, N. Active inference for integrated state-estimation, control, and learning. *arXiv* (2020).

28. Friston, K. J., Mattout, J. & Kilner, J. Action understanding and active inference. *Biological cybernetics* **104**, 137–60 (2011).

29. Oliver, G., Lanillos, P. & Cheng, G. An empirical study of active inference on a humanoid robot. *IEEE Transactions on Cognitive and Developmental Systems* **14**, 462–471 (2021).

30. Friston, K. A theory of cortical responses. *Philosophical transactions of the Royal Society B: Biological sciences* **360**, 815–836 (2005).

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

32. Bogacz, R. A tutorial on the free-energy framework for modelling perception and learning. *Journal of Mathematical Psychology* **76**, 198–211 (2017).

33. Priorelli, M. & Stoianov, I. P. Flexible Intentions: An Active Inference Theory. *Frontiers in Computational Neuroscience* **17**, 1 – 41 (2023). URL https://www.frontiersin.org/articles/10.3389/fncom.2023.1128694.

34. Priorelli, M. & Stoianov, I. P. *Intention Modulation for Multi-Step Tasks in Continuous Time Active Inference* (2022). URL https://link.springer.com/book/9783031287206.

35. Bongard, J., Zykov, V. & Lipson, H. Resilient machines through continuous self-modeling. *Science* **314**, 1118–1121 (2006).

36. Buckley, C. L., Kim, C. S., McGregor, S. & Seth, A. K. The free energy principle for action and perception: A mathematical review. *Journal of Mathematical Psychology* **81**, 55–79 (2017).

37. Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J. & Penny, W. Variational free energy and the Laplace approximation. *NeuroImage* **34**, 220–234 (2007). URL https://www.sciencedirect.com/science/article/pii/S1053811906008822.

38. Friston, K., Parr, T. & de Vries, B. The graphical brain: belief propagation and active inference **1**, 222–241 (2018). URL http://dx.doi.org/10.1162/netn$_{a0}$0083.

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., the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 945539 to GP, the European Research Council under the Grant Agreement No. 820213 to GP, 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.

Supplementary Materials

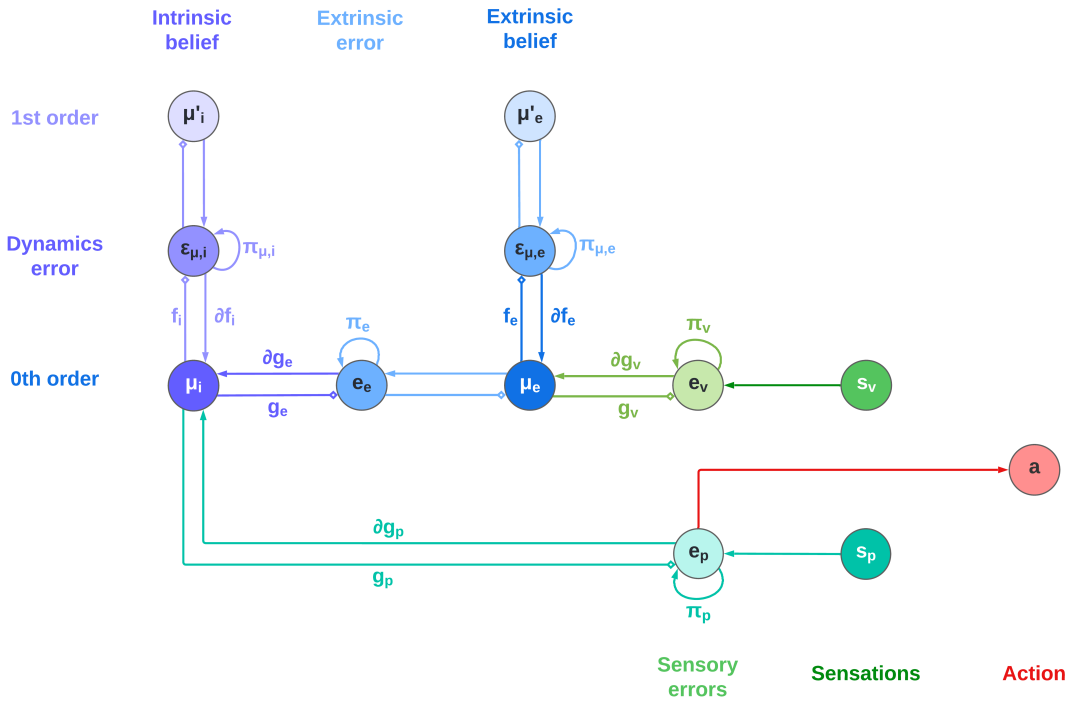


Figure S1: Schematic neural-level implementation of the IE model. The horizontal axis shows the hierarchy between intrinsic beliefs, extrinsic beliefs and sensations. Rather, the vertical axis shows the hierarchy between dynamic levels. The nodes connecting hierarchical levels encode prediction errors. In this schematic, the intrinsic belief directly generates extrinsic predictions of the end effector.

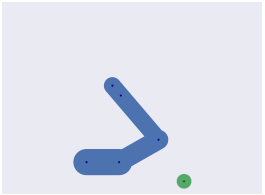
Applications

In this section, we show some further practical applications of the deep kinematic model.

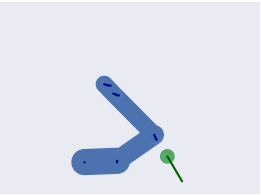
Avoid task

In the Avoid task, the agent has to avoid a moving obstacle with every joint, see Fig. S2. In this task, the agent is only subject to repulsive potentials, specified at every level of the hierarchy.

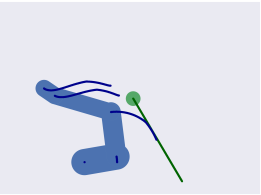
At trial onset (t=0), the obstacle starts in the lower right corner with a random direction.



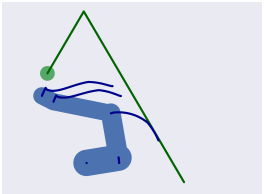
(a) t=0



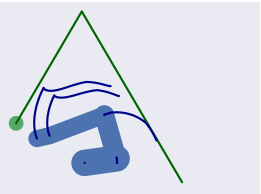
(b) t=150



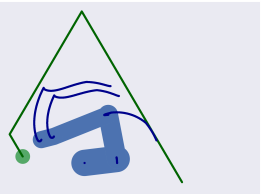
(c) t=500



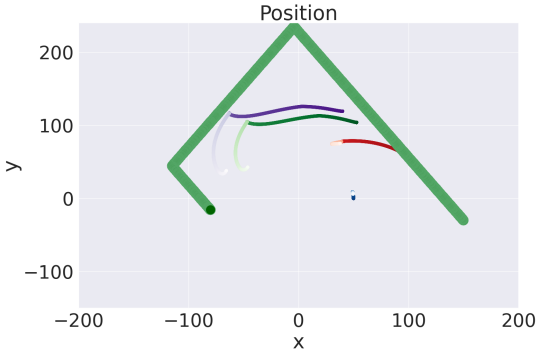
(d) t=1400

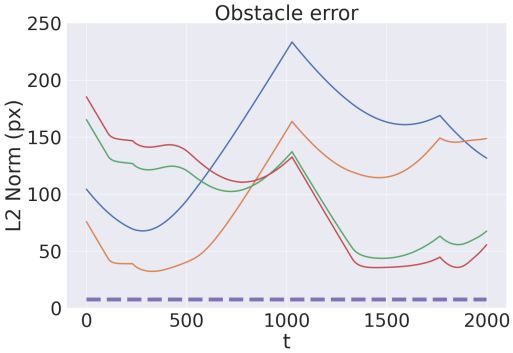


(e) t=1700



(f) t=1900





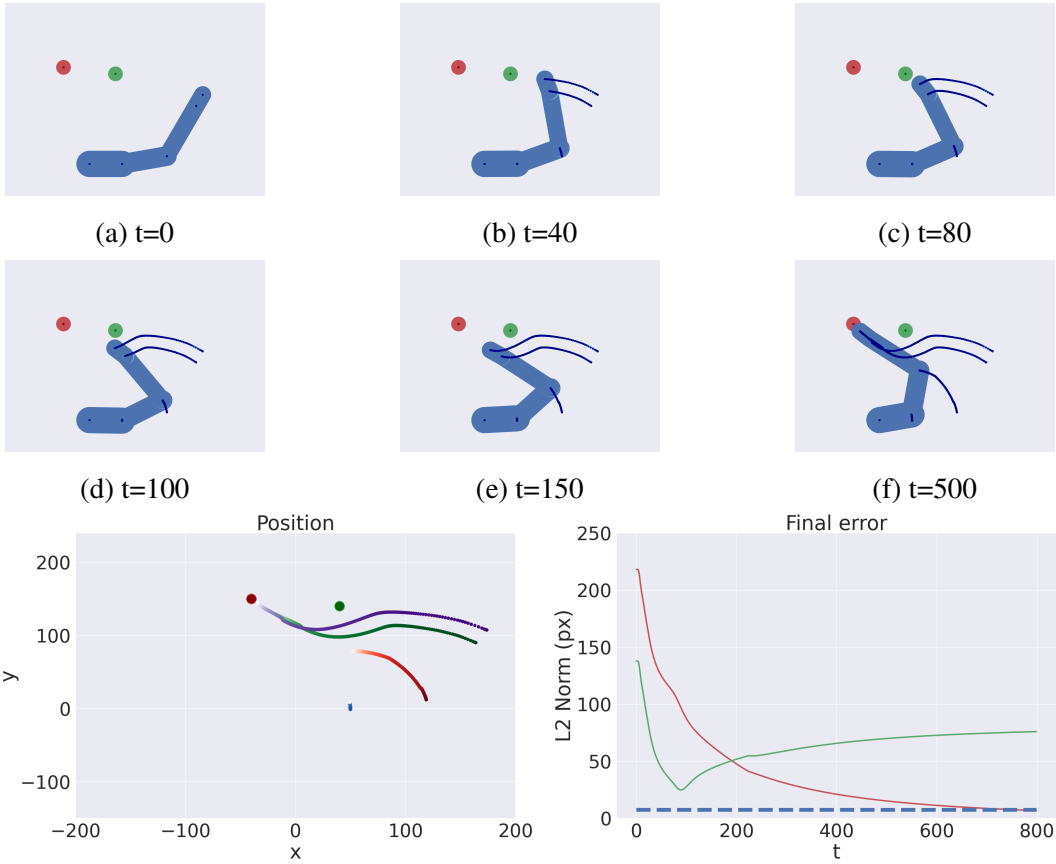
(g) Dynamics of avoid task.

Figure S2: Avoidance of a dynamic obstacle (green circle). (a-f) Sequence of time frames. The arm and obstacle trajectories are shown with blue and green lines, respectively. (g) Trajectory of the arm and the obstacle (left) and distance between the obstacle and each joint (right). The dotted line is the minimum distance required to avoid the obstacle.

At t=150, the agent starts perceiving a repulsive force and moves to avoid the obstacle with the elbow. At t=500, the agent successfully avoids the obstacle. At t=1400, as the obstacle approaches again the arm, the agent perceives another repulsive potential, but now at the end effector position, and avoids again the obstacle in 300 time steps. Finally, at t=1900, the agent performs another slight adjustment. In all cases, it manages to successfully keep at a safe distance, as shown in the right panel of Fig. S2g.

Reach and avoid task

In the reach and avoid task, the agent has to reach a target location with the end effector, while avoiding an obstacle located in the middle of the trajectory, see Fig. S3. Here, the agent is subject to both attractive potentials (only for the end effector) and repulsive potentials (for every joint), encoded as extrinsic beliefs.



(a) t=0 (b) t=40 (c) t=80

(d) t=100 (e) t=150 (f) t=500

(g) Dynamics of the reach and avoid task.

Figure S3: Reach and avoid task. The red and the green circles represent the target location and the obstacle, respectively. (a-f) Sequence of time frames. (g) Arm trajectory (left) and distance between the end effector and the target (red line), and between the end effector and the obstacle (green line) The dotted line represents both the reach distance and the minimum distance to avoid the obstacle.

At t=0, both target and obstacle spawn at random positions. At t=40, the agent moves

only by the attractive force. At t=80, the agent begins perceiving the obstacle and curves its trajectory. At t=150, it has successfully avoided the obstacle, and it moves again only by the attractive potential - although it may still be influenced by repulsive potentials from intermediate joints.

Track and avoid task

The track and avoid task, which is similar the the previous one, except that the target moves, see Fig. S4. The goal of the agent is to track the target for the whole trial, while avoiding the dynamic obstacle.

At t=0, the target and the obstacle spawn in close positions but with different directions. The agent starts to track the target, but at t=1600 it perceives a repulsive force that is stronger than the attractive force from the target. At t=2900, it manages to avoid the obstacle and continues tracking the target.

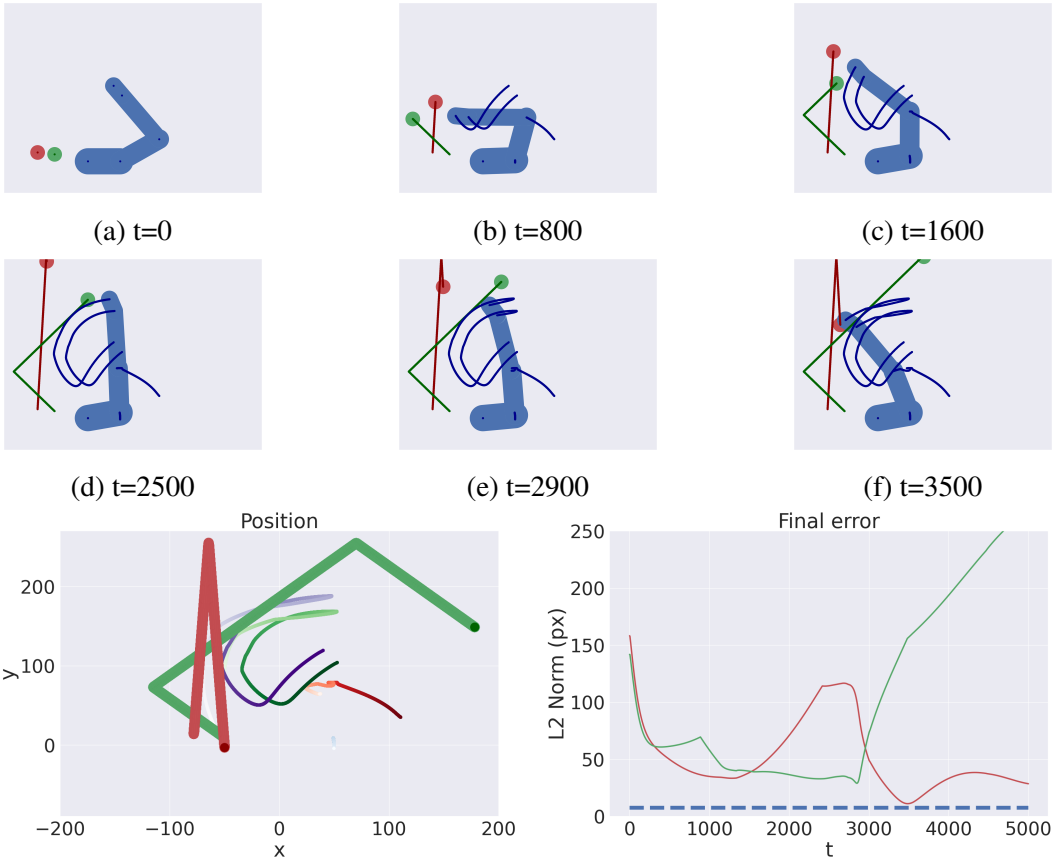
Maintain final orientation task

In this task, the agent has to reach a target location while maintaining a vertical end effector orientation during the whole trajectory (as if it had a glass full of water in the arm and did not want the water to fall), see Fig. S5.

In this task, the agent has two goals. The former consists in setting the end effector position to the target location, as in the previous tasks. The latter sets the end effector absolute orientation to 90 degrees. Different from the previous tasks, we specify every component of the extrinsic belief of the last layer:

$$\boldsymbol{h}_c^{(4)} = \boldsymbol{i}_c^{(4)}(\boldsymbol{\mu}_c^{(4)}) = \begin{bmatrix} x_d & y_d & \phi_d \end{bmatrix} \tag{24}$$

As shown in Fig. S5d, the agent achieves the task successfully, by correctly maintaining the vertical end effector orientation while moving from the initial to the final location.

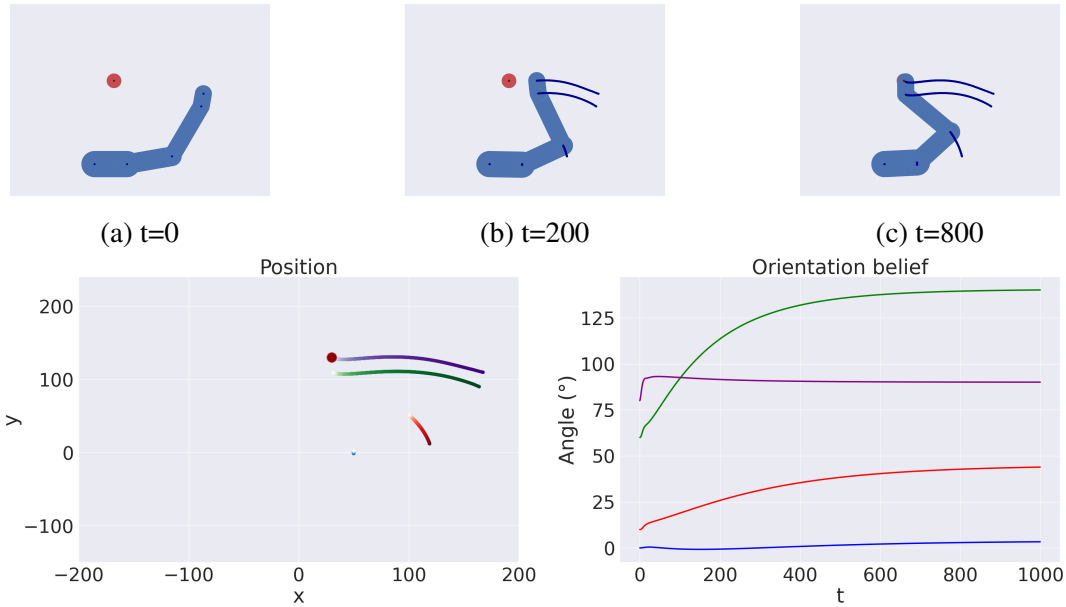


(g) Dynamics of the track and avoid task.

Figure S4: Track and avoid task. (a-f) Sequence of time frames. (g) Arm trajectory (left), distance between the end effector and target (red line), and between the end effector and obstacle (green line). Also see Supplementary Video 1.

Maintain final position task

In this task, the agent has to set the relative joint angle of the shoulder to a specific orientation (here, 100 degrees) while maintaining a fixed end effector position, see Fig. S6. This task requires fulfilling two goals, which are specified at different levels and reference frames. The goal to set the shoulder joint to a specific angle is defined in the intrinsic reference frame, while



(a) t=0 (b) t=200 (c) t=800

(d) Dynamics of the orientation task.

Figure S5: Maintain the same orientation. (a-c) Sequence of time frames. (d) Arm trajectory (left) and absolute orientation of every joint over time (right). The last joint is represented in purple. Also see Supplementary Video 2.

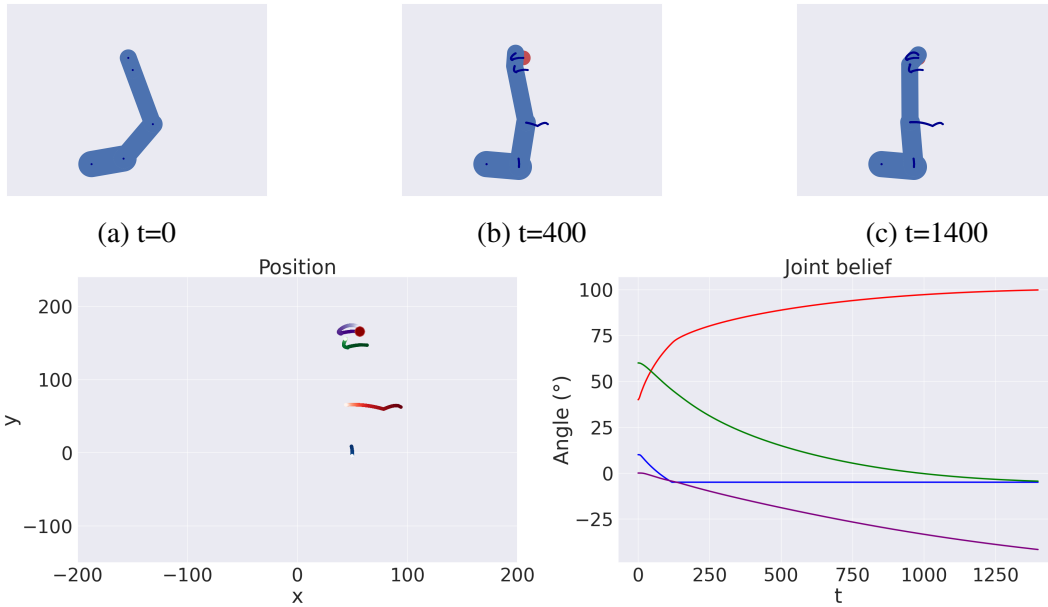
the goal to keep the end effector at a fixed target position is defined in the extrinsic domain.

$$\begin{aligned} \boldsymbol{h}_c^{(4)} &= \boldsymbol{i}_c^{(4)}(\boldsymbol{\mu}_c^{(4)}) = \begin{bmatrix} x_d & y_d & \phi \end{bmatrix} \\ \boldsymbol{h}_j^{(2)} &= \boldsymbol{i}_j^{(2)}(\boldsymbol{\mu}_j^{(2)}) = \begin{bmatrix} \theta_d & l \end{bmatrix} \end{aligned} \tag{25}$$

As shown in Fig. S6d, the agent reaches the desired shoulder angle correctly and rotates the wrist to maintain the same position; yet is slightly moves the end effector during the trajectory, since the joint limits do not allow perfect tracking.

Target pointing task

Sometimes we want to point to a target, rather than reach it (see Fig. S7). While a reaching goal is inappropriate, it is possible to impose an absolute orientation over a segment equal to



(a) t=0 (b) t=400 (c) t=1400

(d) Dynamics of the position task.

Figure S6: Maintain the same position. (a-c) Sequence of time frames. (d) Arm trajectory (left) and relative angles of every joint over time (right). The shoulder joint is represented in red.

the direction of the vector from the segment's extremity to the target:

$$\begin{aligned} \boldsymbol{d} &= \boldsymbol{p}_r - (x^{(4)}, y^{(4)}) \\ \boldsymbol{h}_c^{(4)} &= \boldsymbol{i}_c^{(4)}(\boldsymbol{\mu}_c^{(4)}) = \begin{bmatrix} x & y & \operatorname{atan2}(d_y, d_x) \end{bmatrix} \end{aligned} \tag{26}$$

where $\boldsymbol{p}_r$ is the target position.

Trajectory planning in Cartesian space

Defining attractors in the extrinsic belief is also useful for realizing optimal trajectories in the Cartesian space, as it is possible to define every sort of path, velocity, and acceleration profiles. However, since the extrinsic prediction errors are backpropagated through the inverse kinematic model, the actual movement does not necessarily result in the desired trajectory. Imposing a prior of a linear velocity over the 1st order extrinsic dynamics results in a slightly curved movement, if the update dynamics of both beliefs are similar. However, if the extrinsic update

(a) t=0 (b) t=200

Figure S7: Target pointing. Although the goal is only specified at the level of the last joint, the agent will move freely, until the end effector points towards the target.

is sufficiently slower, the intrinsic belief will follow the desired trajectory, thus allowing to perform linear or circular movements, as shown in Fig. S8.

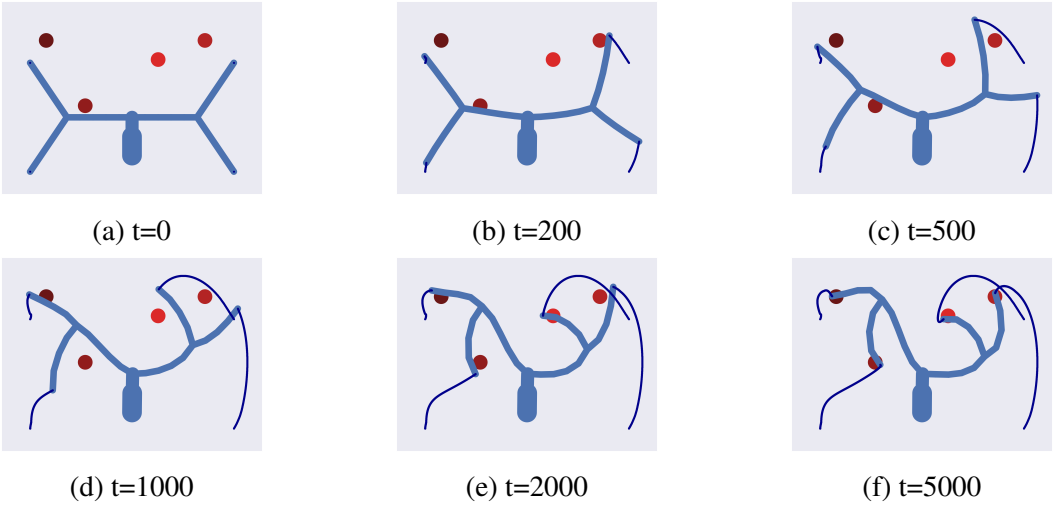
(a) Linear trajectory. (b) Circular trajectory.

Figure S8: Trajectory planning in the Cartesian domain. Slowing down sufficiently the extrinsic dynamics allows to perform linear (left) and circular (right) trajectories. Also see Supplementary Video 3.

Controlling complex kinematic structures

The deep model can deal with complex kinematic structures, such as as the one shown in Fig. S9, composed of 28-DoF and linked in a way as to form a tree of depth 3. The goal of the agent is to reach 4 different target locations with the end of each branch.

This task is more difficult than the previous ones: apart from the much higher number of



(a) t=0 (b) t=200 (c) t=500

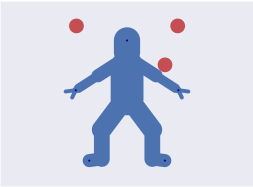
(d) t=1000 (e) t=2000 (f) t=5000

Figure S9: Controlling complex Kinematic Structures. The goal consists in reaching 4 different target locations with a 28-DoF agent. Also see Supplementary Video 4.

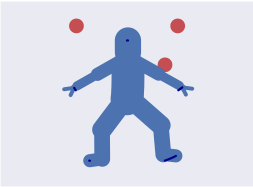
joints, the inference for the high-level layers should balance the contributions from different branches. Yet, the deep model is able to complete the task successfully, despite it requires a considerably higher number of time steps.

Similarly, it is possible to simulate human full-body kinematics. To exemplify this, we performed two tasks, which required controlling a simplified human body composed of 23-DoF, to reach 3 different target locations with the left knee and both hands (Fig. S10) and to avoid a moving obstacle with the whole body (Fig. S11).

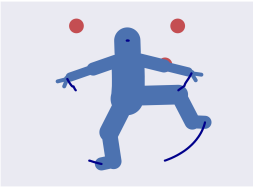
In both simulations, the remaining segments freely move - depending on which ones are constrained - by backpropagation of errors throughout the whole kinematic chain, thus realizing realistic body movements. Other intentions can be imposed if the goal is only to move the desired joints.



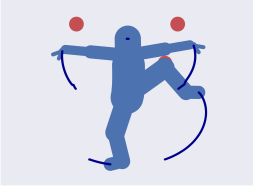
(a) t=0



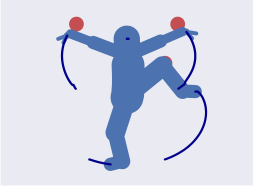
(b) t=100



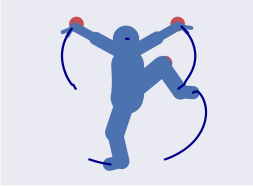
(c) t=500



(d) t=1500

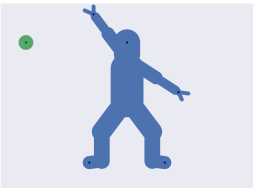


(e) t=2500

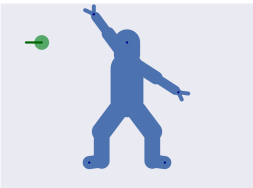


(f) t=3500

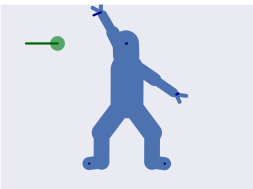
Figure S10: Controlling a simplified human body composed of 23-DoF. The goal is reaching 3 different target locations, with the left knee and the two arms. Also see Supplementary Video 5.



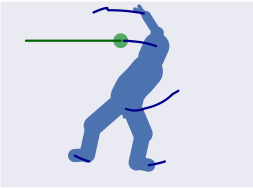
(a) t=0



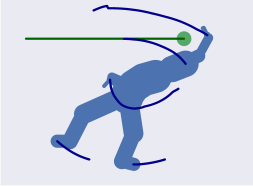
(b) t=500



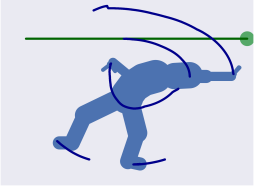
(c) t=1000



(d) t=3000



(e) t=5000



(f) t=7000

Figure S11: Controlling a simplified human body composed of 23-DoF. The task consists in Avoiding a dynamic obstacle with the whole body. Also see Supplementary Video 6.