



World models and predictive coding for cognitive and developmental robotics: frontiers and challenges

Tadahiro Taniguchi, Shingo Murata, Masahiro Suzuki, Dimitri Ognibene, Pablo Lanillos, Emre Ugur, Lorenzo Jamone, Tomoaki Nakamura, Alejandra Ciria, Bruno Lara & Giovanni Pezzulo

To cite this article: Tadahiro Taniguchi, Shingo Murata, Masahiro Suzuki, Dimitri Ognibene, Pablo Lanillos, Emre Ugur, Lorenzo Jamone, Tomoaki Nakamura, Alejandra Ciria, Bruno Lara & Giovanni Pezzulo (2023) World models and predictive coding for cognitive and developmental robotics: frontiers and challenges, *Advanced Robotics*, 37:13, 780-806, DOI: [10.1080/01691864.2023.2225232](https://doi.org/10.1080/01691864.2023.2225232)

To link to this article: <https://doi.org/10.1080/01691864.2023.2225232>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group and The Robotics Society of Japan.



Published online: 26 Jun 2023.



Submit your article to this journal [↗](#)



Article views: 3606










View related articles [↗](#)



View Crossmark data [↗](#)

World models and predictive coding for cognitive and developmental robotics: frontiers and challenges

Tadahiro Taniguchi ^a, Shingo Murata ^b, Masahiro Suzuki^c, Dimitri Ognibene ^{d,e}, Pablo Lanillos^{f,g}, Emre Ugur ^h, Lorenzo Jamoneⁱ, Tomoaki Nakamura^j, Alejandra Ciria ^k, Bruno Lara ^l and Giovanni Pezzulo ^m

^aDepartment of Information Science and Engineering, Ritsumeikan University, Kyoto, Japan; ^bKeio University, Tokyo, Japan; ^cThe University of Tokyo, Tokyo, Japan; ^dUniversità Milano-Bicocca, Milan, Italy; ^eUniversity of Essex, Colchester, UK; ^fDonders Institute for Brain Cognition and Behaviour, Nijmegen, Netherlands; ^gCajal International Neuroscience Center, Spanish National Research Council, Madrid, Spain; ^hBogazici University, Istanbul, Turkey; ⁱQueen Mary University of London, London, UK; ^jThe University of Electro-Communications, Tokyo, Japan; ^kNational Autonomous University of Mexico, Ciudad de Mexico, Mexico; ^lUniversidad Autónoma del Estado de Morelos, Morelos, Mexico; ^mInstitute of Cognitive Sciences and Technologies, National Research Council of Italy, Roma, Italy

ABSTRACT

Creating autonomous robots that can actively explore the environment, acquire knowledge and learn skills continuously is the ultimate achievement envisioned in cognitive and developmental robotics. Importantly, if the aim is to create robots that can continuously develop through interactions with their environment, their learning processes should be based on interactions with their physical and social world in the manner of human learning and cognitive development. Based on this context, in this paper, we focus on the two concepts of *world models* and *predictive coding*. Recently, world models have attracted renewed attention as a topic of considerable interest in artificial intelligence. Cognitive systems learn world models to better predict future sensory observations and optimize their policies, i.e. controllers. Alternatively, in neuroscience, predictive coding proposes that the brain continuously predicts its inputs and adapts to model its own dynamics and control behavior in its environment. Both ideas may be considered as underpinning the cognitive development of robots and humans capable of continual or lifelong learning. Although many studies have been conducted on predictive coding in cognitive robotics and neurorobotics, the relationship between world model-based approaches in AI and predictive coding in robotics has rarely been discussed. Therefore, in this paper, we clarify the definitions, relationships, and status of current research on these topics, as well as missing pieces of world models and predictive coding in conjunction with crucially related concepts such as the free-energy principle and active inference in the context of cognitive and developmental robotics. Furthermore, we outline the frontiers and challenges involved in world models and predictive coding toward the further integration of AI and robotics, as well as the creation of robots with real cognitive and developmental capabilities in the future.

ARTICLE HISTORY

Received 12 January 2023
Revised 10 May 2023
Accepted 22 May 2023

KEYWORDS

World model; cognitive robotics; predictive coding; free-energy principle; active inference

1. Introduction

How can we develop robots that can autonomously explore the environment, acquire knowledge, and learn skills continuously? Creating autonomous cognitive and developmental robots that can co-exist in our society has been considered an ultimate goal of cognitive and developmental robotics and artificial intelligence (AI) since the inception of these fields. Autonomous robots that can develop in the real world and collaborate with us may also be called embodied artificial general intelligence (AGI). The recent success of artificial intelligence depends primarily on large-scale human-annotated data. However, human infants can acquire knowledge and

skills from sensorimotor information through physical interactions with their environment and social interactions with others (e.g. their parents or caregivers). Importantly, the aim is to build robots that can continuously develop through embodied interactions, their learning process must be strongly based on their own sensorimotor experiences. This autonomous learning process that occurs throughout development is also referred to as continual or lifelong learning [1–3], and is considered the foundation for the emergence of both individual and social abilities necessary for robots with adaptive and collaborative capabilities.

Recently, *world models* have attracted renewed attention in artificial intelligence [4–7]. Now, the term ‘world’

CONTACT Tadahiro Taniguchi  taniguchi@ci.ritsumeikan.ac.jp

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group and The Robotics Society of Japan.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

does not indicate the objective world but rather refers to a world understood from a robot's point of view¹. This idea corresponds to that of *Umwelt* proposed by Uexküll [9]. *Umwelt*, literally around-world, meaning environment or surroundings, refers to the self-centered world of an organism perceived through its species-specific sensors². Therefore, notably, the world model is different from the bird's-eye model of the world that was aimed to build in good-old-fashioned AI and criticized later [12]³. A cognitive system learns a world model to predict its future sensory observations better and optimize its policies, also referred to as controllers. Note that although typically the term 'world model' is used to denote the spatiotemporal dynamics of the external environment, it could also equally apply to bodily dynamics (including interoceptive signals from inside the body) and the social environment. This world-model view entails previous ideas and results, such as the effect of behavioral feedback on sensory sampling and perceptual learning [14] and the resulting acquisition of self-centered, yet efficient, representations induced by an active perception strategy on the part of an agent [15].

Predictive coding is another related theory that recently has become more and more influential [16]. It is heavily influenced by Helmholtz's early theories of perception as a process driven by learning, knowledge, and inference [17]. Predictive coding proposes that the brain infers the external causes of sensations by continuously predicting its input through top-down signals and adapts to minimize prediction error [18, 19]. This substantiates the idea that the brain might use an adaptive world model to support perception. The free energy principle (FEP) also proposes a similar vision. It argues that our brain supports both perception (perceptual inference) and action (active inference) using a form of variational Bayesian inference; in particular, using (variational) free energy, it assesses the quality of the prediction and its conformity to prior beliefs [20]. These ideas, which are currently influential in neuroscience and cognitive science, are also used in cognitive and developmental robotics, neuro-robotics [21–23], and artificial intelligence to develop neurodynamics realizing adaptive behaviors and social perception [24].

Although such a learning-driven world model-based approach is promising in cognitive and developmental robotics, the many applications and studies of world models tend to be limited to simulation studies or adopt an offline pretrained world model [25]. Meanwhile, many studies based on predictive coding have been conducted in the field of cognitive robotics and neuro-robotics. However, the relationship between the world model-based approach in AI and the predictive coding-based approach in robotics has rarely been

discussed in an integrated manner. We believe that clarifying the definition, relationship, current state of the art, notable research gaps in work on world models, predictive coding, free-energy principle, and active inference in the context of cognitive and developmental robotics is important for further progress in this field. Based on the current status, we elucidate the frontiers and challenges toward this holy grail in cognitive and developmental robotics.

In this survey paper, we aim to build bridges and clarify the challenges and frontiers of world models and predictive coding in cognitive robotics. The remainder of this paper is structured as follows. Section 2 provides a working definition of each key concept. Section 3 describes prior works related to the concepts and clarifies state of the art. Section 4 describes some notable challenges. Some additional discussion is provided in Section 5, and we conclude the work in Section 6.

2. Working definition

2.1. World model

World models describe the *internal models* of an agent, which encodes how world states evolve, respond to agents' actions, and relate to a given sensory input [4, 26]. The term *world model* dates back to the beginnings of artificial intelligence and robotics [27]. Early research in machine learning studied how an agent could independently acquire and adapt a world model to [28, 29]. Currently, it usually refers to predictive models [30], which are mainly encoded using deep neural networks.

In recent years, advancements in the studies on deep neural networks have enabled self-supervised (or unsupervised) learning⁴ of large-scale world models directly from observations (sensory inputs) [26, 31, 32], and these models have been applied in various areas of artificial intelligence, including reinforcement learning (RL). World models allow agents to perform a sample-efficient prediction of the present state of the world and enable the prediction of future states, which further enables efficient planning (equivalent to model-based reinforcement learning or control). A compact internal representation further enables planning in an efficient low-dimensional space.

The key elements of world models are *prediction* and *inference*⁵. *Prediction* is the probabilistic process of generating the *observation* x given the *state* (or *representation*) z , whereas *inference* is the process of obtaining a state representation z from an observation x in a probabilistic manner. In real-world settings, observations x are large (high-dimensional, e.g. images) and provide only partial information about the world (partial observability). At

the same time, the latent representation z is assumed to represent the internal state of the world. In a static case, these can be summarized as the generative processes of probability distributions as follows.

$$\begin{aligned} \text{Prediction : } & x \sim p(x|z) \\ \text{Inference : } & z \sim q(z|x), \end{aligned} \quad (1)$$

where p is a generative model and q is an inference (or recognition) model⁶. These models are considered parameterized in deep neural networks. When these models are trained simultaneously, generative approaches (e.g. variational autoencoders [38]) are employed [26]. There are also cases where only an inference model is trained, in which case a discriminative approach (e.g. contrastive learning [39]) is used⁷.

In the most common conditions, the state of the environment (and agent body) and the observations evolve over time in response to the agent's actions. In such cases, the state z is often assumed to satisfy Markov conditions. In turn, due to typical sensory limitations such as limited field of view or occlusions, the environment is assumed to follow a partially observable Markov decision process (POMDP) [34]. That is, when the current internal state of the environment is z_t (where the subscript represents a discrete time step), performing an action a_t causes the internal state to transition to z_{t+1} and the corresponding x_{t+1} is observed. In the POMDP case, the prediction and inference models are given as follows.

$$\begin{aligned} \text{Prediction (transition) : } & z_t \sim p(z_t | z_{t-1}, a_{t-1}) \\ \text{Prediction (generation) : } & x_t \sim p(x_t | z_t) \\ \text{Inference : } & z_t \sim q(z_t | x_{1:t}, a_{1:t-1}), \end{aligned} \quad (2)$$

where $x_{1:t}$ denotes the set of observations from the first step (x_1) to step t (x_t). To learn a state space model (SSM) on the time interval 1 to T , a variational approach can be adopted that maximizes the following objective [6, 30, 41]:

$$\begin{aligned} & \log p(x_{1:T} | a_{1:T-1}) \\ & \geq \sum_{t=1}^T \underbrace{\mathbb{E}_{q(z_t | x_{1:t}, a_{1:t-1})} [\log p(x_t | z_t)]}_{\text{(Negative) prediction error}} \\ & \quad - \underbrace{[D_{\text{KL}}[q(z_t | x_{1:t}, a_{1:t-1}) || p(z_t | z_{t-1}, a_{t-1})]]}_{\text{Regularization}} \\ & \equiv \sum_{t=1}^T L_t, \end{aligned} \quad (3)$$

where D_{KL} is the Kullback-Leibler divergence. The first term in Equation (3) represents the prediction error (or reconstruction error) of the observation, and the second

term represents the regularization for the state representation (so that the transition model and the inference model yield the same state representation).

2.2. Predictive coding and the free-energy principle

The original predictive coding model provided by Rao and Ballard [18] was proposed as a model of visual processing in the brain. The model assumes a hierarchically organized neural network, and top-down and bottom-up interactions at each hierarchical level are considered. In the top-down process, higher levels generate predictions about lower-level neural activities, and the lowest level generates sensory predictions. In the bottom-up process, residual errors between the predictions and actual activities (or sensory inputs) are computed and used to correct the originally generated predictions at each level. Predictive coding models learn spatial and temporal statistical regularities at each level for efficient coding and to reduce the redundancy of the predicted activity of lower levels [42, 43]. The main principle behind this hierarchical predictive coding cortical organization in the brain is prediction error minimization (PEM) [19, 44, 45]. This idea based on the principle of PEM has been extended to various cognitive processes, and this framework is usually referred to as predictive processing [16, 46–48]. This approach is being recently used also in machine learning to learn robust generative models of data [49].

The principle of PEM can be situated within a more general principle of free-energy minimization because the amount of variational free energy, the core information measure used in the FEP, can be understood, under simplifying assumptions, as the amount of prediction error [19, 45, 50]. The variational free energy F_t , which is an upper bound on the surprise $-\log p(x_t | x_{1:t-1}, a_{1:t-1})$, is the negative value of the evidence lower bound (ELBO) L_t introduced in Equation (3) as follows.

$$\begin{aligned} -L_t & = F_t \\ & = \underbrace{-\mathbb{E}_{q(z_t | x_{1:t}, a_{1:t-1})} [\log p(x_t | z_t)]}_{\text{Prediction error}} \\ & \quad + \underbrace{\mathbb{E}_{q(z_{t-1} | x_{1:t-1}, a_{1:t-2})} [D_{\text{KL}}[q(z_t | x_{1:t}, a_{1:t-1}) || p(z_t | z_{t-1}, a_{t-1})]]}_{\text{Regularization}} \\ & = \underbrace{D_{\text{KL}}[q(z_t | x_{1:t}, a_{1:t-1}) || p(z_t | x_{1:t}, a_{1:t-1})]}_{\text{Divergence}} \\ & \quad - \underbrace{\log p(x_t | x_{1:t-1}, a_{1:t-1})}_{\text{Evidence}} \\ & \quad \underbrace{\hspace{10em}}_{\text{Surprise}} \end{aligned}$$

$$\geq -\log p(x_t | x_{1:t-1}, a_{1:t-1}). \quad (4)$$

Note that the difference between the evidence and the surprise is only the sign. The evidence and the surprise tend to be used in machine learning studies, e.g. variational inference, and the FEP literature, respectively. From the second line of Equation (4), when observations are assumed to follow a diagonal Gaussian distribution with a fixed variance, minimizing the variational free energy is equivalent to minimizing the sum of mean squared errors and a regularization term.

The FEP is a mathematical formulation of how self-organizing systems, such as biological agents, brains, and cells, are able to maintain an equilibrium with their environment by means of minimizing variational free energy, or the surprise associated with sensations⁸ [44, 51, 52]. In the FEP, different cognitive processes such as perception and action can be understood as different ways to minimize the variational free energy in terms of probabilistic inference called active inference [53–55] as detailed in the next subsection.

2.3. Active inference and exploration

Active inference is a normative framework that derives from the FEP and provides a unifying account for perception, control, and learning in terms of minimization of the variational free energy in the past, present, and future. This unification is important in neuroscience as it reflects neural mechanisms and on a computational level because it offers new perspectives and the possibility of sharing algorithmic solutions between all these functions and transforming them into sophisticated robotic behaviors [23]. For example, perception aims to minimize the variational free energy in the past and present by inferring the latent representations of observed sensory inputs⁹, e.g. when an orange appears in the field of view instead of the apple as currently encoded in the internal representation, the representation state can change toward that of an orange [19]. Conversely, actions try to minimize the variational free energy in the present by actively sampling sensory inputs¹⁰, e.g. by moving the gaze away from the orange toward an apple. In addition to selecting an action in the present, agents can infer a sequence of future actions (or policy) that elicit the most plausible future states [54–58] by considering the minimization of *expected* free energy, as detailed below.

While active inference introduces an important perspective towards an understanding of adaptive and autonomous behaviors, an obvious behavioral imperative, the exploration-exploitation dilemma, seems in conflict with this idea because exploration, i.e. observing an uncertain aspect of the environment, would result

in obtaining an unpredictable outcome [59, 60]. Indeed exploration and active perception have a central role in robot control and learning. Several tasks focus on robots' ability to explore an unknown environment [61–63]. Furthermore, in social contexts, unobservable factors such as others' intentions must be actively considered to allow for efficient human-robot collaboration [64–67].

However, in [20, 55, 68], the authors showed that active inference can easily support exploratory behaviors and that it can provide an elegant formal solution for the exploration-exploitation dilemma.

In fact, we must consider that planning behaviors for an extended period of time requires anticipating future data. More specifically, to infer the best action sequences (policies), one must also predict the future observations they would produce. This is realized in the active inference framework by minimizing the expected free energy over a time interval T . We can express this as the sum of two terms, including (i) the variational information gain term [64, 69–71], or *epistemic* value [55], defined as the expected KL divergence between the distribution of the latent states conditioned on the expected observations $q(z_{t+1:T} | x_{t+1:T}, a_{t:T-1})$ and the prior distribution on the latent states $q(z_{t+1:T} | a_{t:T-1})$ that represents the reduction in uncertainty on the latent states $z_{t+1:T}$ provided by the expected observations $x_{t+1:T}$, and (ii) the extrinsic or *pragmatic* value $\log p(x_{t+1:T} | C)$, where C denotes the agent's preferences, which is a parameter of the distribution $p(x_{t+1:T} | \cdot)$ ¹¹. This results in the following expression¹².

$$G(a_{t:T-1}) = - \underbrace{\mathbb{E}_{q(z_{t+1:T}, x_{t+1:T} | a_{t:T-1})} [D_{\text{KL}}[q(z_{t+1:T} | x_{t+1:T}, a_{t:T-1}) || q(z_{t+1:T} | a_{t:T-1})]]}_{\text{Epistemic value}} - \underbrace{\mathbb{E}_{q(x_{t+1:T} | a_{t:T-1})} [\log p(x_{t+1:T} | C)]}_{\text{Pragmatic value}} \quad (5)$$

where $G(a_{t:T-1})$ is the expected free energy of the future action sequence $a_{t:T-1}$. The epistemic value term favors obtaining observations that disambiguate the world state such as obtaining the address for the best apple shop in town, versus observations that correspond to multiple (aliased) states such as corridors in a mall. Without the factor of variational information gain, asking the address of the shop would not be preferred to any other action that would not immediately result in obtaining an apple. Thus, minimizing expected free energy corresponds to maximizing the sum of epistemic and pragmatic values over an extended period and defines the optimal trade-off between exploration and exploitation. The similarity between the epistemic value term in Equation (5) and the divergence term in Equation (4)

with an inverted sign may be noted. This is due to the different role that observations play in expected free-energy formulation, where they comprise not observed data but expected observations. Finally, the close connection between variational free energy (Equation (4)), expected free energy (Equation (5)), used in this context to define behaviors with exploration capabilities, and ELBO (Equation (3)), used to model learning processes objectives, shows the versatility of this type of formulation, the extension and refinement of which currently a promising field of research that aims to develop an autonomous system with the ability to efficiently acquire and execute complex skills [71]. For a more advanced and detailed presentation, we refer to [20, 55, 71, 72].

Another important framework that considers behaviors as inference is planning or *control as inference* (CaI) [73–77]. The main difference between CaI and active inference is that CaI introduces a binary optimality variable O_t that represents whether an action a_t in state z_t is optimal (or preferred) [77, 78]. If the reward for taking action a_t in state z_t is $r(z_t, a_t)$ ¹³, the conditional distribution of the optimality variable is defined as follows.

$$p(O_t = 1 | z_t, a_t) \equiv \exp(r(z_t, a_t)). \quad (6)$$

Thus, unlike active inference, CaI can introduce the value of the reward at each time explicitly and independently of the observation's generative model¹⁴.

CaI aims to obtain the optimal policy $p(a_t|z_t)$ for inference. If the variational inference is chosen as a solution to the intractability of exact inference (as with active inference), we seek the policy that maximizes the following ELBO¹⁵.

$$\log p(O_{1:T}) \geq \mathbb{E} \left[\prod_{t=1}^T p(a_t | z_t) p(z_t | z_{t-1}, a_{t-1}) \times \left[\sum_{t=1}^T r(z_t, a_t) + H(p(a_t | z_t)) \right] \right], \quad (7)$$

where H represents the entropy. This corresponds to the entropy-regularized expected reward, and reinforcement learning with this as the objective is called entropy-regularized reinforcement learning [80].

3. Prior works

3.1. World models and model-based reinforcement learning in AI and robotics

In this section, we describe world models used in *model-based reinforcement learning* in the context of artificial intelligence and robotics.

In this research field, variational autoencoders (VAE) [38] and recurrent neural networks (RNN) are

frequently used to learn from large-scale spatiotemporal observation data obtained from the external environment. VAE is a type of deep generative model consisting of an encoder that maps inputs to a compact representation called the latent variable and a decoder that generates inputs from the latent variable. These correspond to the inference and prediction in Equation (1) and are trained to maximize the ELBO as shown in Equation (3). RNN is a type of neural network with a recurrent structure that processes temporal data by calculating the next state from the input data at each point in time and the state at the previous point in time. It has been used in various temporal processing fields such as natural language processing and speech recognition.

Time-series world models conditioned on behavior have been studied for policy learning for some time. Schmidhuber proposed learning an agent's policy (utility) via an RNN-based world model obtained by self-supervised learning [28]. Based on this idea, Ha et al. introduced a large-scale world model consisting of a VAE and an RNN that learned directly from observations (time-series images) from the external world [5]. They showed that the policies of agents trained only on this world model, which learns a game environment, can work properly in real game environments. Since this study, research has been conducted on self-supervised learning of models of an environment directly from observations, along with ideas referred to as 'world models'. However, the authors trained spatial compression (VAE) and temporal transitions (RNN) separately; thus, the perspective of learning state representations was not considered.

Subsequently, VAE-based models that simultaneously learn time transitions and spatial compression have been proposed. Kaiser et al. proposed a VAE-based world model with discrete latent variables designed to predict the next frame and reward from the stacked frames of the previous four steps and the current action and showed that model-based reinforcement learning using this model performed adequately in an Atari video game environment with high sample efficiency [83]. One limitation of this model is that it does not include RNNs and cannot account for long-term prediction. Moreover, the measures are learned from the observation space, so the learned representation is not fully exploited. Ke et al. showed that learning long-term transitions using a stochastic RNN-based world model contributes to high performance on tasks that require long-term prediction [84]. All of these models, however, are autoregressive, requiring the generation of a high-dimensional observation space every step for long-term prediction, and are unable to transition within the latent space.

Recently, models that learn transitions in latent space without requiring autoregressive generation have been widely used. Hafner et al. introduced a recurrent state space model (RSSM) that includes RNNs in SSM and showed that it could be used for long-term prediction and model-based reinforcement learning with higher performance than model-free learning [41].

While PlaNet [41], the first study using RSSM, used an existing model-based planning method (cross-entropy method) for planning in the latent space¹⁶, Dreamer [30], a subsequent method, explicitly modeled the policy and value function in neural networks and learned a world model through gradients in an actor-critic framework, resulting in a better performance than PlaNet. This model has been further developed by replacing the latent variables with discrete values, which significantly outperformed model-free performance in Atari game environments (Dreamer V2 [6]), and by using contrastive learning instead of reconstruction, which resulted in higher performance on tasks that were difficult to reconstruct (Dreaming [81], see the left side of Figure 2). They were also combined and compared (Dreaming V2 [86]).

In terms of obtaining a good state representation for control, enforcing explicit constraints on transitions is preferable. For example, NewtonianVAE was able to form PD-controllable state space [87]. However, to develop such a model, what kind of state representation the world model should acquire (as a good representation for control) should be considered, which remains as yet relatively unclear (see Section 4.1 for details).

These world models have been shown to be effective in learning using real robots. Okumura et al. successfully applied the NewtonianVAE to a robot and enabled it to perform a precise socket insertion task [82] (see the right side of Figure 2). Wu et al. showed that Dreamer V2 enabled real robots to perform online learning with very high sample efficiency and performance, which includes a pipeline of acquiring data through interaction with the external world, learning a world model, and controlling the robot using the model [7]. However, all of these results are for a single environment and task, and what kinds of world models should be acquired for robot control in diverse environments and tasks remains unclear.

3.2. Predictive coding and active inference in cognitive and developmental robotics

In recent years, an increasing body of research has considered predictive coding models for perception and action in robotics. Recent comprehensive reviews on active inference and predictive processing in robotics can be found in [21, 23], respectively. These ideas aim

to provide a general mathematical account of behavior. Importantly, they incorporate adaptation and robustness to current methods in cognitive and developmental robotics.

Since the early works of Tani et al. using hierarchically organized RNNs [88], a variety of methods have been proposed to exploit this idea of prediction-error-minimization or propagation. ‘Higher levels’ (internal representation) generate predictions about the dynamics of the ‘lower levels’ up to the sensorimotor level. Prediction errors at the sensorimotor level, given the observations, are then propagated ‘upwards’ in the hierarchy correcting the internal state and thus minimizing the errors. Extensions of Tani’s approach allow multiple time scales [89–92] (see the left side of Figure 3), stochasticity [93, 94] and stochastic latent representations [95]. In particular, a precision-weighting mechanism for the PEM enabled robots to extract stochastic or fluctuating structures of temporal sensorimotor sequences and utilize the extracted structures for their action generation [94]. Interestingly, this mechanism is related to the precision account in psychiatric disorders [96] (especially autism spectrum disorder [97, 98]) and several works have proposed cognitive robot models based on aberrant-precision to model unusual perception and action [99–101].

Aside from hierarchical RNNs, active inference controllers for robotic manipulators [103, 104] and humanoid robots [22, 102] have also been developed based on Lanillos’s initial work on predictive coding adaptive perception and learning [105, 106] for both low-dimensional and high-dimensional inputs (see the right side of Figure 3). These methods have widespread applications such as object manipulation [107], imitation [107], language acquisition [108], social interaction [109] and navigation [110].

Cognitive robots benefit from predictive coding mechanisms to infer others’ actions [111]. The reuse of common circuits for both movement generation and action estimation seems to be a key principle in sensorimotor organization. Recently, the authors of [112] proposed deep modality blending networks (DMBN) designed to create a common latent space from the multi-modal experience of a robot by blending multi-modal signals with a stochastic weighting mechanism. Using a state-of-the-art skill-encoding system referred to as Conditional Neural Movement Primitives (CNMPs) [113], they showed that deep learning could facilitate action recognition and produce structures to sustain anatomical (mirror-like) and effect-based imitation capabilities when combined with a novel modality-blending scheme.

Current state-of-the-art research is focusing on scaling active inference in planning tasks [52] with

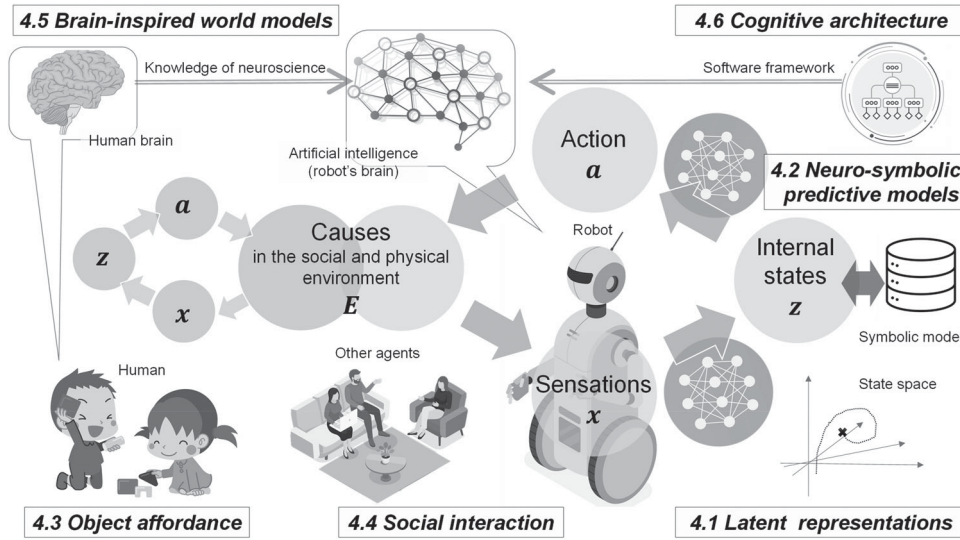


Figure 1. Overview of challenges and relationships between topics described in this survey. A robot, similar to a human, receives *sensations* x , infers *internal states* z , exhibits *actions* a , and affects *causes* E in the social and physical environment. The initial problem is determining the models and architecture that a robot uses to efficiently and effectively learn *latent representations*. An approach to this problem uses a *neuro-symbolic predictive model*, which combines neural network and symbolic models. The notion of *object affordance* highlights the importance of object-centric representation learning and the coupling of action and perception. *Social interaction* with other agents is also an important area of research. Developing artificial intelligence for cognitive and developmental autonomous robots based on knowledge of neuroscience, i.e. *brain-inspired world models*, is promising. Creating *cognitive architecture* and developing and sharing software frameworks for this purpose will also be an important frontier.

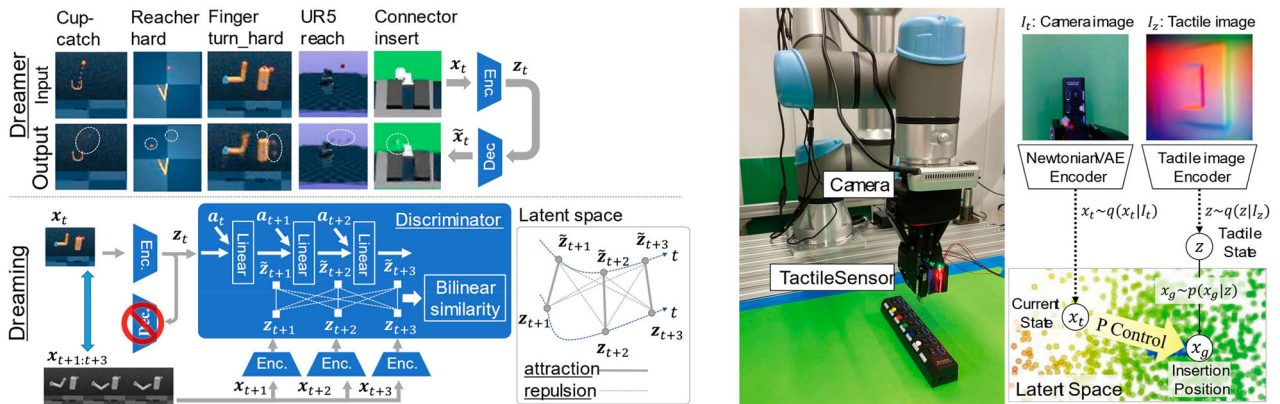


Figure 2. Researches of world models. Left: Dreamer [30] and Dreaming [81]. Right: robot control system using NewtonianVAE [82].

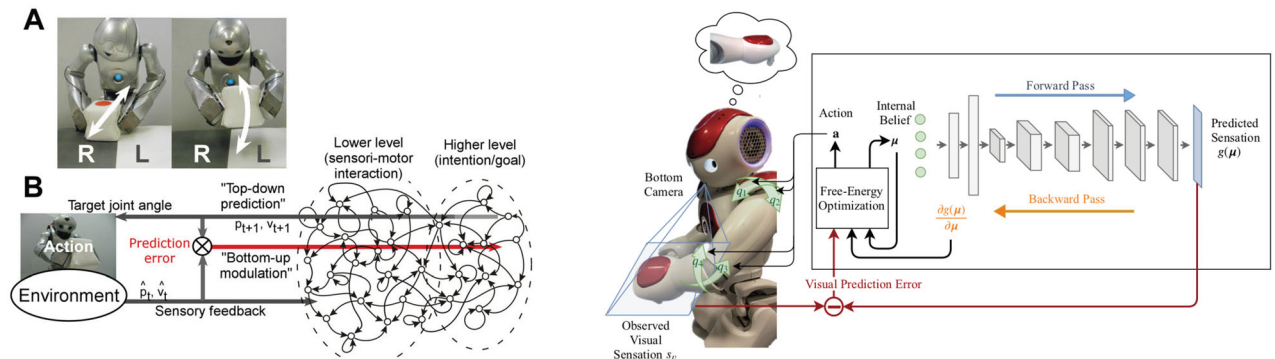


Figure 3. Research on predictive coding and active inference in cognitive robotics. Left: adaptation to environmental changes by prediction error minimization [92]. Right: body perception and action by pixel-based deep active inference [102].

high-dimensional inputs [32, 114] and improving representation learning through multimodal common latent space [112] or introducing structural inductive biases, such as objects [115]. Whilst active inference is a promising framework for robotics [116], current works are still limited to a particular aspect of cognitive and developmental processes. Therefore, in addition to extending the scalability of computational frameworks, continual or lifelong learning for developing abilities from low-level sensorimotor skills to higher-order cognitive functions should also be considered.

4. Frontiers and challenges

4.1. Latent representations for action planning

One of the most important challenges in world-model approaches of any kind is that of efficiently performing planning, in the sense of generating meaningful actions to solve a sequential task [117]. Working in the high-dimensional space of the sensorimotor manifold is very computationally expensive and provides local optima solutions [102]. In fact, current approaches in planning use a compressed encoded representation of the world dynamics, which aids in the process of predicting future states and in action generation [26]. In reinforcement learning, state representation learning is tied to learned tasks to achieve high performance because it depends on the actions needed to obtain the maximum expected reward [30]. However, this sometimes prevents generalization across tasks. Decoupled action-representation world models are an interesting workaround [39]. In deep active inference [23] amortized methods have also been considered [32, 118], in addition to contrastive [119] and iterative amortized inference approaches [115].

However, the key question cannot be narrowed down to that what type of architecture or method should be used. Rather, what type of information should be encoded in the latent representation and how this information is processed must be a key focus so that information is not uncoupled from the sensorimotor process, particularly from motor control, which is a key limitation of existing endeavors in robotics. There has been considerable discussion as to what would comprise an appropriate state representation of a world model; that is, what inductive bias or prior knowledge should be given [120–122]. Here, we list the properties of this prior knowledge we consider important.

- *Low dimensionality.* Observations obtained from the environment are high-dimensional, and compressing this information into a low-dimensional space is

critical for efficient data handling, abstraction, and planning. This approach is the most frequently considered in state representation learning. The challenge is how best to represent observations in a low-dimensional encoding while retaining the necessary task-dependant information. Recent literature focused on generative and discriminative approaches to tackle this.

- *Meaningful abstraction and disentanglement.* Low-dimensional representations should have scene-understanding and task meaning, such as objects [115], locations [123] and temporal events [124]. Representation disentanglement proposes that factors of variation with different semantics should be separated, contributing to the requirements for sufficiency and efficiency in state representation. Object-centric representation learning is related to this hypothesis, [115, 125], in which every observed object is encoded independently.
- *Compositionality.* Although disentanglement aims to separate independent factors, the agent should also acquire their relationships and hierarchy. In the case of object representations, there should also be relations or implication relations among objects. Currently, methods such as those using graph neural networks are being considered, but they do not provide an essential solution. This idea of the compositionality of representation is also relevant to the neuro-symbolic approach.
- *Dynamics prediction.* These three properties are important not only for learning representations in static environments but in dynamic worlds, e.g. they consider a transition model that depends on external factors and agent actions. The best latent representation is one that allows transitions to be easily predictable for given actions. Many recent models use RNNs to learn transitions, which incorporate information on long-term dependence [30, 41]. One way to make transitions more predictable is to incorporate prior knowledge of the physical world (e.g. dynamics following Newton's laws of motion [87]). Furthermore, by learning to separate representations that are not related to control from state representations, representations that are easier to control can be acquired [126].
- *Values are sufficiently encoded.* To perform reinforcement learning on the state representation of the world model, the value of the state representative to the agent must be known. For example, a recurrent state representation learns to predict the reward from the state so that the reward is embedded in the representation [30]. However, because the value of the state changes depending on the task, it

remains unclear whether this hypothesis should be introduced in a world model that should acquire a prediction model that is as task-independent as possible. Alternatively, in active inference approaches, the agent value function cannot be modified, and it is defined by expected free energy. Here, the challenge becomes learning the state preferences and being able to predict the transitions that may yield those preferences.

- *Task-agnostic.* Representations should be informative to solve narrow problems where the agent is trained but also sufficiently general to be reused in tasks with different kinds of variability or new tasks that the agent has never encountered.
- *Fusion of multiple-types multimodal information.* Robots inevitably face a variety of events with their multimodal sensorimotor systems. Observations given to world models are from multiple sources (e.g. social non-social, sensorimotor purely sensorial, and linguistic and non-linguistic). They can have different reliability and volatility and represent various aspects of the world. Therefore, the world models must properly encode the internal representation in a stable and efficient manner.

These are some of the elements that we identified that a latent representation should be fulfilled to provide a smooth connection with real-world interaction and provide power for solving cognitive tasks. Importantly, abstract representation and disentanglement, such as objects or events encoding, may be important to achieve efficient planning, reducing the gap for neuro-symbolic solutions.

In the case of planning actions in the latent space captured by world models, this can be done using forward samples (simulations) from the transition model in that space, especially for simple tasks [127]. For long-term action planning, since the latent space is lower-dimensional than the observation space, guided policy search [128] and evolutionary strategy algorithms such as the covariance matrix adaptation evolution strategy (CMA-ES) [129] and the cross entropy method [130] can be used efficiently [26, 41, 131]. There are also methods to explicitly learn policies for action planning, such as using simulated trajectories to learn policies [132]. In such planning, the transition model of the latent space is not always complete, leading to a problem called the *planning horizon dilemma* [133], i.e. the dynamics must be accurately modeled over long horizons for proper planning, but planning over long sequences often results in poor performance as transition errors accumulate. Therefore, various approaches have been proposed, such as using deep neural networks to extract and utilize only useful

information from simulations with transitions [132]. It is also proposed to learn a latent space that models the inverse dynamics, making good future action planning conditioned on the goal of an arbitrary task [134]. Recent studies have attempted to mitigate this problem by learning policies through the gradient of the value function in simulations over multiple steps of the world model [6, 30], which allows for long-term action planning, such as mining diamonds in Minecraft (where various items need to be acquired before mining diamonds), without human demonstrations [135].

Thus, various studies on learning latent representations and planning actions have been conducted, and significant progress has been made. However, some challenges remain, such as how to connect between the low-dimensional (and hierarchical) encoding and the synchronization with the sensorimotor control.

4.2. Neuro-symbolic predictive models

In this section, we provide an overview of state-of-the-art techniques in which symbols and rules are discovered and used by robots through neuro-symbolic approaches. The term *symbols*, here, refers to manipulative discrete representations used in symbolic AI and cognitive science. The neuro-symbolic approach attempts to integrate conventional symbolic and modern neural network-based AIs.

Both biological and artificial agents benefit from predictive coding mechanisms for reasoning, decision-making, and planning. Predictive forward models are used to generate plans that involve a sequence of actions. For example, chimpanzees are known to generate multi-step plans that include stacking a number of boxes on top of each other, grabbing a long stick, climbing on top of a stack of boxes, and using the stick to reach the object that was initially out of reach [136, 137]. While the underlying cognitive mechanisms for high-level planning remain unknown, different specific brain regions have been shown to become active in inductive and deductive reasoning in humans [138] while predicting the effects of actions [139]. In artificial agents, on the other hand, standard search and planning rely heavily on manually coded or learned state transitions and prediction models [140, Ch. 3–6, 10–11].

The learning of discrete representations of predictive models, specifically dynamic Bayesian networks, to plan goal-directed arm/hand control involved the seminal work conducted by [141]. Ahmetoglu et al. showed that the units generated by slow feature analysis with the lowest eigenvalues resemble symbolic representations that highly correlate with high-level features, which might be

precursors for fully symbolic systems [142]. Konidaris et al. studied methods to discover useful symbols that can be directly utilized in problem and domain definition language (PDDL) for various agent settings [143, 144]. In simulation and the real world, the discovered symbols were directly used as predicates in the action descriptions to generate deterministic and probabilistic symbolic plans. James et al. focused on learning symbols in the ego-centric frame of the agent for transferability to new settings [145]. Ugur et al. combined several machine learning algorithms such as X-means clustering and support vector machine (SVM) classification to enable robots to discover symbols in the continuous perceptual space for PDDL-based manipulation planning [146, 147]. Although the robot discovered the symbols without human intervention, the authors manually encoded the continuous perceptual features. Towards an end-to-end framework, Ahmetoglu et al. developed a deep predictive coding neural architecture that directly used raw camera images and pixel values to discover symbols [148]. In detail, they proposed a deep encoder-decoder network with a binary bottleneck layer designed to take a camera image and an action as input and output the action effects in pixel coordinates. The binary activations in the bottleneck layer encode object symbols that not only depend on the visual input but are also shaped based on action and effect. In other words, the objects that provide the same affordances [149] were automatically grouped together as object symbols. A decision tree was trained to reproduce its decoder function to distill the knowledge represented by the neural network into rules useful for symbolic reasoning. Probabilistic rules were extracted from the effect predictor / neural decoder and encoded in the probabilistic PDDL, which the off-the-shelf AI planners can directly use. In follow-up work, Ahmetoglu introduced a multi-head attention mechanism to enable robots to learn symbols to encode affordances of various objects [150].

Asai et al. implemented a neural framework where a state autoencoder with a discrete bottleneck layer was trained first, and preconditions and effects of actions were learned next [151]. In follow-up work, [152] combined the previous two systems and discovered action preconditions and effects together with visual symbols. These works were realized in visual environments such as 2-D puzzles and achieved visualized plan executions. An important aspect of pure visual neuro-symbolic systems and studies on neuro-symbolic robotics is that in robotics, predictive coding over object symbols takes actions and effects into account in addition to the features of objects and the environment, which facilitates the formation of symbols that are likely to capture object affordances [153–156].

However, in the context of world models, methods to integrate prior symbolic knowledge into VAE and SSM-based world models are still being explored. The bottom-up formation of symbolic representations is closely related to disentanglement and compositionality discussed in Section 4.1. Moreover, leveraging linguistic knowledge in world modeling is a challenge in relation to neuro-symbolic predictive models.

4.3. Affordance perception

Affordance perception has often been discussed independently of world models, but in fact, it is closely related. According to the original definition provided by Gibson [153, 157, 158], an affordance is an action possibility offered to the agent by the environment. A stable surface may afford to be traversed; a stone may afford the possibility of being used as a hammer; a door handle may afford the possibility of being opened. The concept of affordances and affordance perception has then been further analyzed and revisited in psychology, neuroscience, cognitive science, artificial intelligence, and robotics (see [155] for a recent survey). The ability to perceive affordances is crucial for any biological or artificial agent to interact successfully with the environment.

Central to the idea of affordances is that the action possibilities depend on both the agent and the environment; the same environment would offer different action possibilities to different agents, depending on their sensorimotor capabilities. A stable surface affords the possibility of traversal to an agent that is able to locomote and whose body dimensions fit the size of the surface borders; a stone affords the possibility of being used as a hammer to an agent who is able to pick it and who has enough force to lift it; a door handle affords to open to an agent who knows how to open doors, assuming it is well-designed [159]. Therefore, those affordances must be *learned* autonomously by the agent. In fact, the agent must learn how to perceive them. The means by which agents perceive affordances are those of ecological perception, powerfully illustrated by Eleanor Jack Gibson [160–162]: ‘narrowing down from a vast manifold of (perceptual) information to the minimal, optimal information that specifies the affordance of an event, object, or layout’ [162, p.284]. The agent must learn what minimal information is to be picked; this happens both through evolution and development, leveraging the sensorimotor exploration of the environment by physical interaction.

Interestingly, while exploring the action possibilities, the agent can learn the effects of those actions as well. This is crucial for biological agents and turns out to be

extremely useful for artificial systems as well. In fact, most computational models of affordances in robotics rely on representations that include not only the action but also the effects (or in other terms, the goal) of the action [111, 113, 149, 163–173]; therefore, the perceived possibilities for actions (and for achieving certain effects) can be used for action planning, leading to problem-solving [174]. Such comprehensive models of affordances are, in fact, world models; they are internal models of how the world behaves ‘in the eyes’ of the learning agent, and they can be used by the agent to make predictions about how the world will change if certain actions are performed. Therefore, it is not surprising that the computational techniques used for learning affordance models often overlap with those used for learning world models [175]. It is worth noting that, in world model approaches, a robot only receives raw sensory information and needs to extract the relevant semantics from such data flow; therefore, to successfully integrate affordance perception in these systems, the challenges of meaningful abstraction/disentanglement and object-centric representation learning, described in Section 4.1, are particularly relevant.

4.4. Social interaction

Robots’ ‘worlds’ consist not only of physical objects but also of social entities, i.e. people who give robots social guidance and try to cooperate with them. World models should enable agents to understand and predict the social dynamics of human behavior by identifying latent variables such as intentions and emotions. This understanding can allow agents to collaborate with humans to influence and shape social outcomes.

Safe human-robot interaction and efficient collaboration are some of the major research goals in robotics and have important practical applications [176–182]. Theory of mind is the cognitive ability to associate beliefs, intentions, or mental states with other agents, or in other words, to predict the internal state of another agent’s world model and to understand its activities and context [183]. This ability is an essential component of human interaction [184–186] and has attracted attention in robotics [187].

World models can play an important role in mutual understanding under challenging conditions for agents’ perceptual systems, where interactions or just observed behaviors are complex and lead to conflicting interpretations. A prototypical example is the false belief task, where an agent with sensory limitations fails to detect environmental changes and behaves according to its outdated beliefs [188, 189]. It is also crucial when different levels of knowledge and expertise induce different

representations of a domain, as well as different points of view, which may induce different support strategies [190] or even conflictual interactions [191]. For example, the perspective of an automotive mechanic and that of an ordinary car user differ considerably, so collaboration may be difficult if one cannot properly infer others’ knowledge level.

The recent progress in machine learning methods has resulted in substantial improvement in action recognition methodologies [67, 192–194]. However, most work has focused on using shallow and purely perceptual representations to interpret observed activity. This approach often results in limited flexibility in terms of context, task, and observed actors. These methods require significant retraining time and large amounts of data that are difficult to collect and transfer to relatively similar conditions [195].

Alternatively, approaches such as goal recognition as planning or inverse planning [177, 196–198] have shown significant advantages in terms of flexibility, using a model of the environment to understand the activities of others by computing plans that would lead to the observed actions. These approaches have demonstrated the advantages of incorporating a world model into intention recognition. Several works have extended this approach. The problem of dealing with behaviors generated under partial observability, which may require inferring both the plan and the beliefs, the mental state [187, 189], of the observed actor, was studied with both classical planning [199] and Bayesian approaches [188, 200]. The impact of missing observations for the observer agent has also been analyzed [199]. A further step has been proposed by active methods for activity recognition [64, 201, 202] that use the same world model both to interpret others’ actions as well as selecting actions that would improve the recognition process, e.g. by giving access to the most informative observations [65] and allow the completion of a joint task [190]. While even the initial formulations of this approach were computationally aware [196], their efficiency is often affected by the length of the observed behavior and the environment complexity, resulting in methods that can seldom be applied online on a robot. Several models proposed a pre-compile approach that transformed the world into a form that would allow efficient plan recognition [195]. The adoption of hierarchical world model representations has also been applied to limit computational and modeling costs of the process [203–205]. Precomputed and robust local plans, in the form of the same motor controllers that the robot uses to perform its own actions, have also been adopted to allow active perception for action recognition and prediction on humanoid robots [64, 183].

One of the main problems with inverse planning is that implementations often assume that observed behaviors are rational or, in terms of reinforcement learning and planning, optimal. In addition to being expensive to compute, optimal behaviors may differ substantially from the suboptimal or bounded rational behaviors that human collaborators may perform [206, 207]. When a human partner makes a mistake that can lead to failure, the difference between optimal and sub-optimal behaviors can lead to misinterpretations in a context where the human partner is more in need of assistance from the robot. Some approaches attempt to address this condition by allowing for random noise in action selection [208] or by using simplified punishment strategies for sub-optimality, e.g. increased expected plan length [199]. However, human errors and behavioral suboptimality often depend on environmental features that determine the complexity of the underlying cognitive processes [206] or induce habits [207]. An interesting approach to plan recognition that accounts for human cognitive limitations uses a symbolic world model to implement online Bayesian inference about the interplay between incomplete plan execution and online planning or replanning when incomplete plans lead to failure [209].

The additional flexibility provided by world models in social interaction skills is likely relevant beyond activity recognition. It is easy to imagine that shallow purely supervised models may be limited in terms of perspective-taking and the ability to reason based on the world structure may help to adapt to partners with different sensory systems [187, 210–212]. Similarly, world models are likely to help with imitation learning by dealing with embodiment mismatch between the observed actor and the learner [213]. Finally, physical cooperation [214, 215] and signaling [216, 217] would also be more flexible when integrating world and partner models in the equation, for example, to account for the trust of the human cooperator towards the robot [218]. Finally, a world model may also be learned through socially rich experiences and sources of information (e.g. imitation [213] or verbal instructions) in addition to the results of autonomous exploration. However, developing a robust, efficient, and flexible enough representation may prove to be one of the main challenges in this effort.

4.5. Brain-inspired world models

In cognitive science, it has long been postulated that the brain learns small-scale models of the world and uses these models for various cognitive functions, such as perception, planning, and imagination [16, 46–48,

219]. For example, theories of perception-as-inference described perception as an inferential process, which works by ‘inverting’ a generative model of how the percepts are generated [220, 221]. As discussed above, these ideas (and others) have been recently formalized under the label of the *Bayesian Brain* [222] and extended by *Active Inference* from the domain of perception to other domains, such as action planning and interoception [20].

In parallel, there have been many attempts to describe mathematically and to assess the neuronal underpinnings of world models and of inference processes empirically (e.g. [223, 224]). One question that has received a great deal of attention is how the brain might encode internal world models in the neuronal substrate. Given that the brain models are often assumed to be probabilistic, various formal schemes have been proposed that describe plausible neuronal implementations of probabilistic variables and of Bayesian inference over these variables, such as, for example, probabilistic population codes [225] and sampling schemes [226]. These attempts show that (probabilistic, generative) world models could be at least potentially implemented in neuronal substrate [227, 228] – and updated after statistical learning [229] – but the specific scheme(s) that the brain might use for this remain to be fully assessed.

Another relevant question is what algorithms the brain might use to perform inference over world models. A strong candidate in neuroscience is *predictive coding* [18, 19]. Several studies have aimed to validate its key empirical predictions, showing that under the appropriate conditions, it is possible to observe predictions [230], prediction errors [231] and other signatures of inference in brain signals [232] and that neural activity in lower visual areas in the absence of bottom-up inputs could be explained by the top-down, feedback dynamics postulated by predictive coding [233]. These and other studies (see [234, 235] for recent reviews) lend some support for predictive coding, but the theory remains under development.

At yet another level, one may ask what the systems-level architecture that supports world models and whether different parts of the brain might model different aspects of the world is. Anatomical considerations suggest that the brain is not a monolithic entity but rather is composed of several areas and networks [236]; however, the extent to which these areas or networks are modularized and how they exactly influence each other is heavily discussed [237]. One interesting consideration is that cortical brain areas in humans and monkeys appear to be organized along principal gradients (defined by functional connectivity); in one of these gradients, heteromodal areas (e.g. prefrontal cortex) are placed at the top, and unimodal areas (e.g. primary visual

area) at the bottom, recapitulating the structure of a putative hierarchical generative model [238]. Another interesting consideration is that there seems to be a ‘division of labor’ between brain pathways that perform complementary computations, such as the two visual pathways for processing ‘what’ and ‘where’ information [239]. These anatomical and functional separations might be potentially interpreted as useful *factorizations* of the brain generative models. A whole-brain probabilistic generative model (WB-PGM) approach attempts to build a cognitive architecture for cognitive and developmental robots integrating probabilistic generative model (PGM)-based modules referring comprehensive knowledge of human and animal brain architectures and their anatomy [240].

The above studies indicate that at a general level, both neuroscience and machine learning / AI conceive world models and inference in similar ways. However, at a more detailed level, there might be profound differences between the ways these two disciplines use the same concepts. Predictive coding and other biological schemes proposed in neuroscience exploit top-down dynamics (and recurrences) in ways that are rarely used in machine learning. Furthermore, brain information processing is heavily based on spontaneous brain dynamics, which are largely absent in machine learning systems; see [241–243] for a detailed discussion of putative computational roles of spontaneous dynamics. Furthermore, biological brains are energy efficient, whereas the cognitive, metabolic, and other costs of learning and inference [244] are rarely considered in the design of contemporary AI systems. Moreover, it is plausible to assume that different parts of the brain might be specialized (or might have different *inductive biases*) to process different statistical regularities, rendering them able to learn and model (for example) slower or faster dynamics of the visual scenes, one’s own body, the actions of other agents, or extended temporal events [245]. It is also worth highlighting here that, although prediction errors have a central role in learning, the brain might also use other forms of statistical learning that are not error-driven [246]. It remains to be understood how to best endow our more advanced machine learning systems with the ability of the brain to perform (apparently) specialized computations but also orchestrate them coherently. Finally, it is important to remember that the brain is an evolved system, and our more advanced cognitive abilities are grounded in (the neuronal mechanisms supporting) simpler sensorimotor skills [247, 248]. Trying to develop advanced cognitive systems without the necessary requirements for embodied interaction and ‘phylogenetic refinement’ might lead to solutions that differ completely from how the brain works – or that fail altogether.

4.6. Cognitive architectures

Truly cognitive and developmental robots, i.e. embodied AGI, that behave autonomously and flexibly in the real environment would have a wide range of sensors and exhibit multiple functions. That requires a large-scale world model that deals with multimodal sensory observations and multilayered state representations. Considering the discussion in Section 4.5, such world models may be factorized in a proper manner from engineering and biological viewpoints. To realize embodied AGIs, further frameworks and architectures to factorize a total world model into cognitive modules and to integrate individual cognitive capabilities into a cognitive system are required. The idea is related to *cognitive architectures*, which have been studied in cognitive science, artificial intelligence, and robotics [249, 250].

In cognitive science, cognitive functionalities such as memory, perception, and decision-making are implemented as modules in the cognitive architectures studied, and the specific task can be solved by activating these modules coordinately. ACT-R [251] and Soar [252] are representatives of cognitive architectures. It has been shown that the model implemented by ACT-R can explain the time to solve the task by humans, and activation patterns of the brain can be predicted by activation patterns of the modules [253]. Furthermore, Soar has been used for controlling robots [254] and learning games [255]. However, complex machine learning methods that have rapidly advanced in a decade are not introduced yet. Sigma [256, 257] is a newer cognitive architecture that introduces the generative flow graph, a generalized probabilistic graphical model. Therefore, the model can be implemented using probabilistic programming techniques [258–260]. Furthermore, the concept of the standard model of the mind is discussed through a synthesis across these three cognitive architectures [237]. Particularly, cognitive architectures based on first principles, e.g. with a general computation scheme, such as free energy minimization [19], are especially attractive. The architecture for social cognition has also been proposed [261]. The authors point out that these architectures explained above are incomplete in dealing with the social aspect of cognition and describe the elements of architecture for social cognition. Clarion [262] is another cognitive architecture based on dual process theory [263]. In this architecture, each subsystem is composed of explicit and implicit processes, and it is shown that the interaction between implicit-explicit processes can explain psychological phenomena.

In robotics, several types of cognitive architecture have been proposed. One of them is ArmarX [264], which

has three layers, including a middleware layer, a robot framework layer, and an application layer. This three-layered structure simplifies the development of robotics software easier. (Neuro-)Serket [265, 266] is another approach to integrating cognitive modules¹⁷. In (Neuro-)Serket, modules are described by the (deep) PGM and trained mutually by exchanging messages between modules. To make it easy to develop large-scale models, the modules in Neuro-Serket are weakly connected through the Serket interface. (Neuro-)SERKET is closely related to the world model-based approach because SERKET requires each module to be a PGM, i.e. a model based on prediction and inference as Equation (1), and integrate modules into a large PGM. This architecture does not provide any restrictions regarding the functionalities of modules. Therefore, it has high flexibility but brings high dimensional design space at the same time. To reduce the large degree of freedom in the design space, a brain-inspired approach, WBA-PGM, was proposed [240]. In this approach, a cognitive model was constructed by connecting PGM-based modules utilizing knowledge from neuroscience. By referring to the brain studies, WBA-PGM constrains the function of modules and their connection and reduces the design space of the cognitive model.

There are two crucial requirements for cognitive architecture for cognitive and developmental robots, which can be used along with the approach based on world models and predictive coding. The first is the engineering aspect which is seen in (Neuro-)Serket and ArmarX. The scale of cognitive models that enables the robots to behave flexibly in the real environment is very large, and many modules must be connected and work collaboratively. Furthermore, the model needs to introduce machine learning techniques that are not only existing as well as those will be developed in rapid progress. The development of such a model would require a massive engineering effort, and this is considered a notable obstacle to realizing such robots. Therefore, architecture is needed to simplify development. Another requirement is that of the scientific aspect seen in ACT-R, Soar, WBA-PGM, and Clarion. Developing AGI, which is human-like intelligence, referring to the knowledge regarding humans obtained in cognitive science and neuroscience, can accelerate its development. However, meeting these two aspects completely is very challenging. All machine learning techniques and module connections might necessarily be not reasonable from the point of view of cognitive science and neuroscience. On the other hand, entire humans are not understood yet. Therefore, finding common ground between engineering and science aspects and developing a novel cognitive architecture is a current challenge. Developing a large-scale cognitive

architecture and overcoming the problems described in the previous subsections is also a challenge.

5. Discussion

As we described, world models and predictive coding are promising approaches in cognitive and developmental robotics. We summarize the references we mentioned in Sections 3 and 4 in Tables 1 and 2, respectively.

Before closing this paper, we will mention some remaining issues which have not been addressed in the main body sufficiently.

5.1. Language and world models

Umwelts, i.e. worlds from first-person views, of biological systems are not monolithic but have some sort of structure. Notably, language and symbolic systems have syntactic structures. The interaction between high-level cognitive capabilities, e.g. language and reasoning, and low-level cognitive capabilities, e.g. perception and action, is essential in world modeling. Recently, large-scale language models (LLMs) have been replacing many natural language processing methods [270, 271], including reasoning tasks, which have been conducted solely by symbolic AI by the end of 2010s [272, 273]. Recently, the use of LLMs in robotics has been attempted, e.g. [274]. It is clear that language learning and understanding by robots is itself a frontier [275]. To leverage the symbolic knowledge in LLMs, integration of LLMs and world models will be an important challenge.

This shift from models of artificial symbols in conventional AI to models of natural language, i.e. a human symbol system, is resonating with the discussion in symbol emergence in cognitive and developmental systems [276–279]. An important topic is then considering not only the integration of human language into robots' world models in a top-down manner but also the bottom-up formation of symbol systems, including language in relation to world models.

Table 1. Summary of references in Section 3.

World models and model-based reinforcement learning in AI and robotics	
Early work in world models	[5, 28]
Autoregressive world models	[83, 84]
World models with latent space transitions	[6, 30, 41, 81, 85, 86]
World models in robotics	[7, 82, 87]
Predictive coding and active inference in cognitive and developmental robotics	
Reviews on active inference and predictive processing in robotics	[21, 23, 116]
Predictive coding models and extensions	[88–95]
Applications of predictive coding and active inference	[22, 99–113, 267]
Scaling and improvements of active inference	[32, 52, 112, 114, 115]

Table 2. Summary of references in Section 4.

Latent representations for action planning	
Deep active inference and related methods	[23, 32, 115, 118, 119]
Discussion on state representation learning and its inductive bias	[115, 120–122]
Meaningful abstraction and disentanglement	[115, 123–125]
Dynamics prediction	[30, 41, 87, 126]
Action planning methods in latent representation	[6, 26, 30, 41, 127, 131–135]
Neuro-symbolic predictive models	
Biological and artificial predictive coding mechanisms	[136, 138, 139]
Continuous representation and symbol learning	[142, 150, 164, 268, 269]
Discrete representation and symbol learning	[141, 143–148]
Neural frameworks and integration	[115, 148, 151, 152, 156]
Affordance perception	
Ecological perception	[153, 157, 158, 160–162]
Computational models	[137, 149, 154, 155, 159]
Affordance models in robotics	[111, 113, 163–175]
Social interaction	
Human-robot collaboration and interaction	[176, 177, 179–182, 195]
Theory of mind and mental states	[177, 183–189]
Activity recognition and machine learning methods	[65, 67, 192–195]
Goal recognition and planning approaches	[177, 187–189, 196–201, 203, 205]
Active methods for activity recognition	[64–66, 190, 201, 202]
Computational efficiency and world model representations	[64, 183, 196, 204, 209]
World models in social interaction skills	[187, 190, 205, 209–218]
Brain-inspired world models	
Theories of world models and perception-as-inference	[4, 16, 20, 46–48, 219–222]
Putative neuronal implementations of world models	[223–229]
Predictive coding in the brain	[18, 19, 230–235]
Brain architecture potentially supporting world models	[236–240]
Bridging the gap between neuroscience and machine learning	[241–243, 245–248]
Cognitive architectures	
Cognitive architectures overview and perspective	[237, 249, 250]
Traditional cognitive architectures	[251–255]
Modern cognitive architectures and techniques	[19, 256–263]
Robotics cognitive architectures	[240, 264–266]

5.2. Policy representations

How should the policies of robots be represented? Conventionally, policies are described as feedback controllers $\pi(z_t, a_t) = p(a_t | z_t)$ in reinforcement and imitation learning. Even though one direction of world model approaches is to explore task agnostic representations (Section 4.1), the decomposition of world modeling and policy learning can be controversial. In a conventional approach of world models, policies ($\pi = p(a_t | z_t)$ or

$a_{t:T}$) and world models ($p(z_{t+1} | z_t, a_t)$ and $p(o_t | z_t)$) are decoupled. In contrast, a series of studies about predictive coding in neuro-robotics have been intentionally entangling policies and world models and making robots directly learn $p(o_{t+1}, a_{t+1} | o_{1:t}, a_{1:t})$ and exhibiting many successful results in robotics, e.g. [24, 280, 281]. As the notion of affordance also suggests, actions and perceptions are not independent and entangled, generally. The question ‘to what extent should we decouple world model and policy representations?’ should be investigated.

5.3. From artificial cognition to human cognition

Cognitive and developmental robotics are also constructive approaches to human developmental cognition. Not only learning from neuro-, cognitive and developmental sciences but also provide with scientific feedback to them is also an important mission. Building a virtuous circle between studies on human and artificial studies is a challenge.

The constructive approach may give us a novel approach to scientific and philosophical hard problems like self-awareness [282] and consciousness. The relationship between the multimodal world model and global workspace theory was suggested [283]. Extending the discussion between world models and consciousness using robots may be an exciting challenge. Moreover, the relationship between predictive coding and emotion is worth exploring to build emotional robots and understand the emotions of biological systems [284–286].

5.4. Software frameworks for implementation

To accelerate the studies on world models and predictive coding in robotics, the development framework for cognitive and developmental robotics is crucially important. In robotics, not only AI ‘software’ frameworks but also middle-ware is important. Recently, ROS has been widely used in the robotics community for bridging hardware and AI software layers. Developing and sharing such software frameworks as a community will be important, e.g. [287]. Moreover, the world model involves many types of knowledge, and the knowledge can be used for achieving multiple functions via active inference. The software framework should allow the world model to efficiently organize the knowledge and to perform (cross-modal) active inference. A great initiative is the discrete state-space active inference python library [288]. However, it can only be used for toy examples due to scalability issues and to be useful in cognitive and developmental robotics. It needs further development. For instance, the support for high-dimensional input observations and the possibility of combining discrete and continuous action

and state representations are something that has been addressed in robotic approaches [23].

5.5. Data-efficient and autonomous learning

A generalist agent called GATO was developed based on Transformers and shown to be able to solve various tasks with one neural network [289]. Although the approach is superficially different, the approach is really related to the world models and the predictive coding approach. However, the learning system is hugely data-hungry. It is very questionable if the model can be regarded as a model of human intelligence. Moreover, to train the generalist agent, researchers need to prepare a large dataset and simulation environment. Human children can autonomously explore their environment and acquire data through active exploration. Moreover, they use heuristics and biases in their developmental process. Learning and considering the human developmental process will give us the inspiration to build real generalist agents. Developing a data-efficient autonomous learning architecture with world models and predictive coding at its core is the key to a truly cognitive and developmental system.

Online adaptation to the non-stationary environment is one of the keys to realizing autonomous robots. Predictive coding can enable robots to adapt to the non-stationary environment by minimizing prediction errors. Although the world model itself can continuously learn the environmental dynamics in an online manner, the policy learning will suffer from the non-stationarity if the world model-based approach relies on RL, because RL assumes the stationary environment. However, if the world model-based approach uses other planning methods that are robust to non-stationarity, the world model may also allow the robots to adapt to the non-stationary environment. This is also an important issue to investigate.

5.6. Emergence of behaviors

Should an agent have a completely internal model of its world? Lastly, we raise fundamental speculation about the world model-based approach. Behaviors are not externalization of internally designed trajectories but something to emerge through the interaction between the body and the environment. For example, it has been proven by passive walking machines that the behavior of walking emerges only from the interaction between the body and the environment, without any computation by the brain [290]. About three decades ago, Brooks famously advocated the physical grounding hypothesis together with subsumption architecture, saying the world

is its own best model [291]. The robots behaved smoothly and flexibly without any explicit world models. This is also referred to as morphological computation, which means the body itself implicitly processes information dynamically [292, 293]. Soft robotics emphasizes these points nowadays. Combining the viewpoints of the emergence of behaviors with complex physical dynamics and the world model-based approach is another important challenge.

6. Conclusion

In this survey paper, we have aimed to clarify the frontiers and challenges of world models and predictive coding in cognitive and developmental robotics. Creating an autonomous robot that can actively explore the real environment, acquire knowledge, and learn skills continuously is the ultimate goal of cognitive and developmental robotics. To make the robot continuously develop through active exploration, the robot's learning process should be based on sensorimotor information obtained through physical and social interactions with the physical and social environment. Following the motivation, this paper reviewed studies related to world models and predictive coding in cognitive and developmental robotics and related AI studies. We clarified the definition of world model and predictive coding in robotics, in conjunction with those of FEP and active inference, and discussed the relationship between them. We also introduced state-of-the-art and research gaps of studies on world models and predictive in robotics. We described six frontiers and challenges, i.e. latent representations for action planning, neuro-symbolic predictive models, affordance perception, social interaction, brain-inspired world models, and cognitive architecture. Through the survey and clarification of challenges, we provided future directions for developing cognitive and developmental robots based on world models and predictive coding.

Notes

1. This viewpoint may be called a robot's subjective point of view of the world. Philosophically, however, whether robots can have a 'subjective' point of view remains controversial [8]. Therefore, we describe this point of view simply as 'a robot's point of view'.
2. Importantly, the relationship between Umwelt and world modeling was suggested in semiotics. Sebeok pointed out that the closest equivalent of Umwelt in English is 'model' [10]. An Umwelt is created and constructed through a functional cycle, which includes (1) anticipation of a perceptual cue, (2) perception, (3) working out a relation between the perception and action (either simply executing a habit or using representation, or modeling anew), and (4) action (operation) [11].

3. Also, notably, the world-model approach is different from behavior-based robotics [13], which does not learn world models.
4. Self-supervised learning is a type of unsupervised learning that aims to accomplish a task by learning to predict or classify any part of unsupervised data from any other part. In contrast to self-supervised learning, unsupervised learning includes clustering.
5. The use of these terms is sometimes incongruent in the literature on statistics and machine learning; the word *inference* is also used to describe prediction or substituted by other concepts, such as *encoding* and *decoding* in the literature on (variational) autoencoders.
6. In classical formulations of control theory [33], AI [34], and probabilistic robotics [35], the inference step is performed by exactly inverting the prediction probability p using the Bayes theorem. However, this poses several computational challenges and is often intractable. For dealing with such complexity, sampling-based approximate inference can be performed using Monte Carlo methods [36, 37]. In the context of world models and predictive coding, variational Bayes approaches are often preferred [20]. Variational Bayes involves the definition of an approximate inference distribution \hat{q} . Amortized inference allows us to approximate the $q(x)$ using a neural network, which is called an inference network, and obtain an inference model $q(x|a)$ [38].
7. However, recently, studies have also shown that contrastive learning can be interpreted as a generative approach [40]. Therefore, the boundary between generative and discriminative approaches is, to a certain extent, blurred.
8. In contrast to states of surprise, free-energy can be measured because it is a function of sensory input and the inferred state [45]
9. This perception can be regarded as a variational Bayesian version of the original predictive coding that employs a maximum a posteriori (MAP) estimation [44].
10. More specifically, the prediction error in the variational free energy in Equation (4) is the difference between sensory predictions and sensory inputs. Perceptions help make sensory predictions more closely match sensory inputs, while actions help make sensory inputs more closely match sensory predictions.
11. Usually, C is given as a hyperparameter. However, it is possible to train it, for example, by maximizing log-likelihood.
12. Note that in the literature on active inference, a sequence of actions $a_{t:T-1}$ is referred to as a policy π . This policy is different from a policy in reinforcement learning, where it represents a statistical mapping from states to actions ($\pi(a_t|s_t)$). Using this notation of the policy π and mean-field approximation, the general formulation of the expected free energy can be described by the following more practical formulation. $G(\pi) = -\sum_{\tau=t+1}^T \mathbb{E}_{q(x_\tau|\pi)} [D_{\text{KL}}[q(z_\tau|x_\tau, \pi)||q(z_\tau|\pi)]] - \mathbb{E}_{q(x_\tau|\pi)} [\log p(x_\tau|C)]$, where the time step $\tau > t$ used here is a future time step.
13. We assume $r(z_t, a_t) \leq 0$. If $M \equiv \max_{z_t, a_t} (r(z_t, a_t)) > 0$, we can redefine the reward function $r(z_t, a_t) \leftarrow r(z_t, a_t) - M$ to satisfy the inequality without losing generality. See [79] for more details.
14. Therefore, unlike active inference, CaI does not require the assumption of POMDP.

15. Here, MDP is assumed, i.e. state z_t is an observed variable rather than a latent variable; if POMDP is assumed, inference and generative models for observations are added to this ELBO.
16. PlaNet was extended to be uncertainty-aware on the basis of Bayesian inference [85].
17. Neuro-SERKET is an updated version of SERKET.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was partially supported by JST Moonshot R&D Program, Grant Number JPMJMS2033 and JPMJMS2011, JST PRESTO Grant Number JPMJPR22C9, the BAGEP Award of the Science Academy, TUBITAK ARDEB 1001 program (project number: 120E274), the European Union's Horizon 2020 Framework Programme for Research and Innovation under Specific Grant Agreements No. 945539 (Human Brain Project SGA3), No. 952215 (TAILOR), and No. 824153 (POTION), the European Research Council under the Grant Agreement No. 820213 (ThinkAhead), the Human Brain Project Specific Grant Agreement 3 grant (ID 643945539, for the SPIKEREFERENCE project), Deepself project under the Priority Programme "The Active Self" (SPP 2134), the project 'COURAGE - A social media companion safeguarding and educating students' (No. 95563 and No. 9B145), the Volkswagen Foundation inside the initiative Artificial Intelligence and the Society of the Future, and the PNRR MUR project PE0000013-FAIR".

Notes on contributors

Tadahiro Taniguchi received his M.E. and Ph.D. degrees from Kyoto University, in 2003 and 2006, respectively. From 2005 to 2008, he was a Japan Society for the Promotion of Science Research Fellow in the same university. From 2008 to 2010, he was an Assistant Professor at the Department of Human and Computer Intelligence, Ritsumeikan University. From 2010 to 2017, he was an Associate Professor in the same department. From 2015 to 2016, he was a Visiting Associate Professor at the Department of Electrical and Electronic Engineering, Imperial College London. Since 2017, he is a Professor at the Department of Information Science and Engineering, Ritsumeikan University, and a Visiting General Chief Scientist at Panasonic (Holdings) Corporation. He has been engaged in research on machine learning, emergent systems, cognitive robotics, and symbol emergence in robotics.

Shingo Murata received his B.E., M.E., and D.E. degrees from Waseda University in 2011, 2013, and 2016, respectively. From 2016 to 2018, he served as a Research Associate in the Department of Modern Mechanical Engineering at Waseda University. Following that, from 2018 to 2020, he held the position of Assistant Professor at the Principles of Informatics Research Division, National Institute of Informatics (NII). During this period, he was also appointed as an Assistant Professor at the Department of Informatics, The Graduate University for Advanced Studies (SOKENDAI). Since 2020, he has been a

Senior Assistant Professor in the Department of Electronics and Electrical Engineering at Keio University. His current research interests include cognitive robotics, robot learning, and computational psychiatry.

Masahiro Suzuki is a project assistant professor in the Graduate School of Engineering at the University of Tokyo. He was formerly a project researcher at the University of Tokyo from 2018 to 2020. He received his PhD from the University of Tokyo in 2018 and his MS degree from Hokkaido University in 2015. His research interests are in deep generative models, multimodal learning, and transfer learning.

Dimitri Ognibene graduated in Information Engineering at the University of Palermo in 2004 and his PhD in Robotics in 2009 from University of Genoa. He was research associate at ISTC-CNR in Rome since 2005 till 2010. In 2010 he joined the Personal Robotics Lab of Imperial college. In 2014 he moved to at Centre for Robotics Research, Kings' College London. In 2015 he was awarded a Marie Curie Actions COFUND fellowship and joined the AI and ML group of UPF, Barcelona. In 2017 he became lecturer in Computer Science and Artificial Intelligence at the University of Essex, where he still holds a visiting position. He is Associate Professor of Human Technology Interaction at the University of Milan-Bicocca, Italy since 2020. His main interest lies in understanding how social agents with limited sensory and computational resources can adapt to complex and uncertain environments. To this end it develops both neural and Bayesian models and applies them both in the physical field, e.g. robots, and virtual, e.g. social media, settings.

Pablo Lanillos received the Ph.D. degree in computer science engineering from the Department of Computer Engineering and Automatic Control, Complutense University of Madrid, Spain, in 2013. He is principal investigator at the Cajal International Neuroscience Center, Spanish National Research Council, and the Donders Institute for Cognition in the Netherlands. He was awarded the Marie Skłodowska-Curie H2020 Grant at the Technical University of Munich and gained tenure as Assistant Professor at Radboud University Nijmegen. His team develops neuroscience-inspired artificial intelligence algorithms for achieving human-like perception and action in robots and he currently coordinates three international projects where the leitmotif is to transform our understanding of human cognition into technologies of the future.

Emre Ugur is an Associate Professor in Dept. of Computer Engineering, Bogazici University, and the chair of the Cognitive Science MA Program. He received his PhD degree in Computer Engineering from Middle East Technical University, Turkey. He worked as a research scientist at ATR, Japan (2009–2013); visited Osaka University as a specially appointed Assist. & Assoc. Professor (2015 & 2016); and worked as a senior researcher in University of Innsbruck (2013–2016). He is interested in robotics, robot learning, and cognitive robotics.

Lorenzo Jamone is a Senior Lecturer in Robotics and Director of the CRISP team at the School of Engineering and Materials Science (SEMS) of the Queen Mary University of London (QMUL), in the UK. The CRISP team is part of the ARQ (Advanced Robotics at Queen Mary) group. Since 2018 he is a Turing Fellow at The Alan Turing Institute. He received the

MS (Hons) degree in computer engineering from the University of Genoa, Genoa, Italy, in 2006, and the PhD degree in humanoid technologies from the University of Genoa and the Italian Institute of Technology, in 2010. He was an Associate Researcher with the Takanishi Laboratory, Waseda University, Tokyo, Japan, from 2010 to 2012, and with VisLab, Instituto Superior Técnico, Lisbon, Portugal, from 2013 to 2016. His main research interest is Cognitive Robotics, with more than 100 publications (h-index 25) on object manipulation, tactile sensing, affordance perception, sensorimotor learning and control.

Tomoaki Nakamura received his BE, ME, and Dr. of Eng. degrees from the University of Electro-Communications in 2007, 2009, and 2011. From April 2011 to March 2012, He was a research fellow of the Japan Society for the Promotion of Science. In 2013, he worked for Honda Research Institute Japan Co., Ltd. From April 2014 to March 2018, he was an Assistant Professor at the Department of Mechanical Engineering and Intelligent Systems, the University of Electro-Communications. Since April 2019, he has been an Associate Professor at the same department. His research interests are intelligent robotics and machine learning.

Alejandra Ciria obtained her Ph.D. in Experimental Psychology from the Universidad Nacional Autónoma de México (UNAM). In 2019 she did a postdoc in the Adaptive Systems group at the Humboldt- Universität zu Berlin funded by the Alexander von Humboldt foundation. Since 2020, she is an associate professor in the department of Cognitive and Behavioural Sciences at the Faculty of Psychology at the UNAM and the head of the Embodied Cognition Lab. She is a Diverse Intelligences Summer Institute (DISI) fellow. Her research focuses on the study of perceptual, predictive, attentional, and emotional mechanisms under the embodied cognition framework.

Bruno Lara received a PhD in Mechatronics from King's College London. He did a postdoc at the TheoriLabor in the University of Jena, working on evolutionary robotics, and then a postdoc in the Max Planck Institute for Psychology Research in Munich, focusing on research on Cognitive Robotics. He is a Professor at the Universidad Autónoma del Estado de Morelos (UAEM) and the Head of the Cognitive Robotics Lab at the Science Research Centre since 2005. In 2011 and then again in 2019, he did long research stays in the Adaptive Systems group at the Humboldt- Universität zu Berlin. He is an Alexander von Humboldt Fellow. His research interests include internal models, prediction and predictive processing in the framework of cognitive robotics.

Giovanni Pezzulo received his MSc from the University of Pisa and his Ph.D. degree from the University of Rome La Sapienza. He is Research Director at the Institute of Cognitive Sciences and Technologies, National Research Council of Italy, in Rome. He uses a combination of theoretical, computational and empirical methods to study cognitive processing in humans and other animals, with a focus on goal-directed behavior, decision-making, prediction, and planning. His research was funded by the European Research Council, the EU's Horizon 2020, the Human Frontier Science Program, and other agencies. Website: <https://sites.google.com/site/giovannipezzulo/>.

ORCID

Tadahiro Taniguchi  <http://orcid.org/0000-0002-5682-2076>
 Shingo Murata  <http://orcid.org/0000-0002-0790-1915>
 Dimitri Ognibene  <http://orcid.org/0000-0002-9454-680X>
 Emre Ugur  <http://orcid.org/0000-0001-9597-2731>
 Alejandra Ciria  <http://orcid.org/0000-0002-9216-9297>
 Bruno Lara  <http://orcid.org/0000-0002-9844-6435>
 Giovanni Pezzulo  <http://orcid.org/0000-0001-6813-8282>

References

- [1] Lungarella M, Metta G, Pfeifer R, et al. Developmental robotics: a survey. *Conn Sci.* 2003;15(4):151–190. doi: [10.1080/095440090310001655110](https://doi.org/10.1080/095440090310001655110)
- [2] Lesort T, Lomonaco V, Stoian A, et al. Continual learning for robotics: definition, framework, learning strategies, opportunities and challenges. *Inform Fusion.* 2020;58:52–68. doi: [10.1016/j.inffus.2019.12.004](https://doi.org/10.1016/j.inffus.2019.12.004)
- [3] De Lange M, Aljundi R, Masana M, et al. A continual learning survey: defying forgetting in classification tasks. *IEEE Trans Pattern Anal Mach Intell.* 2021;44(7):3366–3385. doi: [10.1109/tpami.2021.3057446](https://doi.org/10.1109/tpami.2021.3057446)
- [4] Friston K, Moran RJ, Nagai Y, et al. World model learning and inference. *Neural Netw.* 2021;144:573–590. doi: [10.1016/j.neunet.2021.09.011](https://doi.org/10.1016/j.neunet.2021.09.011)
- [5] Ha D, Schmidhuber J. World models. preprint [arXiv:1803.10122](https://arxiv.org/abs/1803.10122). 2018.
- [6] Hafner D, Lillicrap T, Norouzi M, et al. Mastering Atari with discrete world models. preprint [arXiv:201002193](https://arxiv.org/abs/201002193). 2020.
- [7] Wu P, Escontrela A, Hafner D, et al. Daydreamer: world models for physical robot learning. preprint [arXiv:220614176](https://arxiv.org/abs/220614176). 2022.
- [8] Kiverstein J. Could a robot have a subjective point of view? *J Conscious Stud.* 2007;14(7):127–139.
- [9] Von Uexküll J. A stroll through the worlds of animals and men: a picture book of invisible worlds. *Semiotica.* 1992;89(4):319–391. doi: [10.1515/semi.1992.89.4.319](https://doi.org/10.1515/semi.1992.89.4.319)
- [10] Sebeok TA. Biosemiotics: its roots, proliferation, and prospects. *Semiotica.* 2001;2001(134):61–78. doi: [10.1515/semi.2001.014](https://doi.org/10.1515/semi.2001.014)
- [11] Kull K. Umwelt and modelling. In: *The routledge companion to semiotics*. New York: Routledge; 2009. p. 65–78.
- [12] Brooks R. Intelligence without representation. *Artif Intell.* 1991;47(1-3):139–159. doi: [10.1016/0004-3702\(91\)90053-M](https://doi.org/10.1016/0004-3702(91)90053-M)
- [13] Arkin RC. Behavior-based robotics. Cambridge, MA: MIT Press; 1998.
- [14] Verschure PF, Voegtlin T, Douglas RJ. Environmentally mediated synergy between perception and behaviour in mobile robots. *Nature.* 2003;425(6958):620–624. doi: [10.1038/nature02024](https://doi.org/10.1038/nature02024)
- [15] Ognibene D, Baldassare G. Ecological active vision: four bioinspired principles to integrate bottom-up and adaptive top-down attention tested with a simple camera-arm robot. *IEEE Trans Auton Ment Dev.* 2014;7(1):3–25. doi: [10.1109/TAMD.2014.2341351](https://doi.org/10.1109/TAMD.2014.2341351)
- [16] Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral Brain Sci.* 2013;36(3):181–204. doi: [10.1017/S0140525X12000477](https://doi.org/10.1017/S0140525X12000477)
- [17] Helmholtz HV. *Handbuch der physiologischen optik*. Leipzig: Leopold Voss; 1867. (Allgemeine Encyclopädie der Physik; vol. III).
- [18] Rao RP, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci.* 1999;2(1):79–87. doi: [10.1038/4580](https://doi.org/10.1038/4580)
- [19] Friston K. A theory of cortical responses. *Philos Trans Royal Soc B: Biol Sci.* 2005;360(1456):815–836. doi: [10.1098/rstb.2005.1622](https://doi.org/10.1098/rstb.2005.1622)
- [20] Parr T, Pezzulo G, Friston KJ. Active inference: the free energy principle in mind, brain, and behavior. Cambridge, MA: MIT Press; 2022.
- [21] Ciria A, Schillaci G, Pezzulo G, et al. Predictive processing in cognitive robotics: a review. *Neural Comput.* 2021;33(5):1402–1432. doi: [10.1162/neco_a_01383](https://doi.org/10.1162/neco_a_01383)
- [22] Oliver G, Lanillos P, Cheng G. An empirical study of active inference on a humanoid robot. *IEEE Trans Cognitve Develop Syst.* 2021;14:462–471. doi: [10.1109/TCD S.2021.3049907](https://doi.org/10.1109/TCD S.2021.3049907)
- [23] Lanillos P, Meo C, Pezzato C, et al. Active inference in robotics and artificial agents: survey and challenges. preprint [arXiv:211201871](https://arxiv.org/abs/211201871). 2021.
- [24] Tani J. Exploring robotic minds: actions, symbols, and consciousness as self-organizing dynamic phenomena. New York: Oxford University Press; 2016.
- [25] Ibarz J, Tan J, Finn C, et al. How to train your robot with deep reinforcement learning: lessons we have learned. *Int J Rob Res.* 2021;40(4-5):698–721. doi: [10.1177/0278364920987859](https://doi.org/10.1177/0278364920987859)
- [26] Ha D, Schmidhuber J. Recurrent world models facilitate policy evolution. *Adv Neural Inf Process Syst.* 2018;31:2450–2462.
- [27] Nilsson NJ. Shakey the robot. 1984. Tech Rep.
- [28] Schmidhuber J. Making the world differentiable: on using self-supervised fully recurrent neural networks for dynamic reinforcement learning and planning in non-stationary environments. München: Institut für Informatik; 1990.
- [29] Sutton RS. Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In: *Machine Learning Proceedings 1990*. Austin, TX: Elsevier; 1990. p. 216–224.
- [30] Hafner D, Lillicrap T, Ba J, et al. Dream to control: learning behaviors by latent imagination. preprint [arXiv:191201603](https://arxiv.org/abs/191201603). 2019.
- [31] Levine S, Finn C, Darrell T, et al. End-to-end training of deep visuomotor policies. *J Machine Learn Res.* 2016;17(1):1334–1373.
- [32] van der Himst O, Lanillos P. Deep active inference for partially observable mdps. In: Verbelen T, Lanillos P, Buckley CL, De Boom C, editors. *Active inference*. Ghent: Springer International Publishing; 2020. p. 61–71.
- [33] Kalman RE. A new approach to linear filtering and prediction problems. *J Basic Eng (Transactions of the American Society of Mechanical Engineers).* 1960;82:35–45. doi: [10.1115/1.3662552](https://doi.org/10.1115/1.3662552)
- [34] Cassandra AR, Kaelbling LP, Littman ML. Acting optimally in partially observable stochastic domains. In:

- AAAI. Vol. 94. Seattle, WA: The MIT Press; 1994. p. 1023–1028.
- [35] Thrun S, Burgard W, Fox D. Probabilistic robotics. Cambridge, MA: MIT Press; 2005.
- [36] Chen Z. Bayesian filtering: from Kalman filters to particle filters, and beyond. *Statistics*. 2003;182(1):1–69. doi: [10.1080/02331880309257](https://doi.org/10.1080/02331880309257)
- [37] Thrun S. Particle filters in robotics. In: Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence. Alberta: Morgan Kaufmann Publishers Inc. 2002. p. 511–518.
- [38] Kingma DP, Welling M. Auto-encoding variational bayes. preprint [arXiv:13126114](https://arxiv.org/abs/1312.6114). 2013.
- [39] Laskin M, Srinivas A, Abbeel P. CURL: contrastive unsupervised representations for reinforcement learning. In: International Conference on Machine Learning (ICML). Virtual, MLR Press; 2020. PMLR 119. p. 5639–5650.
- [40] Nakamura H, Okada M, Taniguchi T. Self-supervised representation learning as multimodal variational inference. preprint [arXiv:220311437](https://arxiv.org/abs/220311437). 2022.
- [41] Hafner D, Lillicrap T, Fischer I, et al. Learning latent dynamics for planning from pixels. In: International Conference on Machine Learning (ICML). Long Beach, CA: MLR Press; 2019. p. 2555–2565.
- [42] Huang Y, Rao RP. Predictive coding. *Wiley Interdisciplinary Rev Cognitive Sci*. 2011;2(5):580–593. doi: [10.1002/wcs.v2.5](https://doi.org/10.1002/wcs.v2.5)
- [43] Hogendoorn H, Burkitt AN. Predictive coding with neural transmission delays: a real-time temporal alignment hypothesis. *Eneuro*. 2019;6(2):1–12. doi: [10.1523/ENEURO.0412-18.2019](https://doi.org/10.1523/ENEURO.0412-18.2019)
- [44] Friston K, Kiebel S. Predictive coding under the free-energy principle. *Philosoph Trans Royal Soc B: Biological Sci*. 2009;364(1521):1211–1221. doi: [10.1098/rstb.2008.0300](https://doi.org/10.1098/rstb.2008.0300)
- [45] Friston K. The free-energy principle: a unified brain theory? *Nat Rev Neurosci*. 2010;11(2):127–138. doi: [10.1038/nrn2787](https://doi.org/10.1038/nrn2787)
- [46] Clark A. Surfing uncertainty: prediction, action, and the embodied mind. Oxford: Oxford University Press; 2015.
- [47] Hohwy J. New directions in predictive processing. *Mind Lang*. 2020;35(2):209–223. doi: [10.1111/mila.v35.2](https://doi.org/10.1111/mila.v35.2)
- [48] Hohwy J. The predictive mind. Oxford: Oxford University Press; 2013.
- [49] Ororbia A, Kifer D. The neural coding framework for learning generative models. *Nat Commun*. 2022;13(1):1–14. doi: [10.1038/s41467-022-29632-7](https://doi.org/10.1038/s41467-022-29632-7)
- [50] Friston K, Kilner J, Harrison L. A free energy principle for the brain. *J Physiology-Paris*. 2006;100(1-3):70–87. doi: [10.1016/j.jphysparis.2006.10.001](https://doi.org/10.1016/j.jphysparis.2006.10.001)
- [51] Friston K, Mattout J, Kilner J. Action understanding and active inference. *Biol Cybern*. 2011;104(1):137–160. doi: [10.1007/s00422-011-0424-z](https://doi.org/10.1007/s00422-011-0424-z)
- [52] Buckley C, Kim C, McGregor S, et al. The free energy principle for action and perception: a mathematical review. *J Math Psychol*. 2017;81:55–79. doi: [10.1016/j.jmp.2017.09.004](https://doi.org/10.1016/j.jmp.2017.09.004)
- [53] Friston KJ, Daunizeau J, Kilner J, et al. Action and behavior: a free-energy formulation. *Biol Cybern*. 2010;102(3):227–260. doi: [10.1007/s00422-010-0364-z](https://doi.org/10.1007/s00422-010-0364-z)
- [54] Friston K, Adams RA, Perrinet L, et al. Perceptions as hypotheses: saccades as experiments. *Front Psychol*. 2012;3:151. doi: [10.3389/fpsyg.2012.00151](https://doi.org/10.3389/fpsyg.2012.00151)
- [55] Friston K, Rigoli F, Ognibene D, et al. Active inference and epistemic value. *Cogn Neurosci*. 2015;6(4):187–214. doi: [10.1080/17588928.2015.1020053](https://doi.org/10.1080/17588928.2015.1020053)
- [56] Kruglanski A, Jasko K, Friston K. All thinking is ‘wishful’ thinking. *Trends Cogn Sci (Regul. Ed.)*. 2020;24(6):413–424. doi: [10.1016/j.tics.2020.03.004](https://doi.org/10.1016/j.tics.2020.03.004)
- [57] Friston K, Samothrakis S, Montague R. Active inference and agency: optimal control without cost functions. *Biol Cybern*. 2012;106(8-9):523–541. doi: [10.1007/s00422-012-0512-8](https://doi.org/10.1007/s00422-012-0512-8)
- [58] Parr T, Friston K, Pezzulo G. Generative models for sequential dynamics in active inference. *Cogn Neurodyn*. 2023:1–14. doi: [10.1007/s11571-023-09963-x](https://doi.org/10.1007/s11571-023-09963-x)
- [59] Friston K, Thornton C, Clark A. Free-energy minimization and the dark-room problem. *Front Psychol*. 2012;3:130. doi: [10.3389/fpsyg.2012.00130](https://doi.org/10.3389/fpsyg.2012.00130)
- [60] Sun Z, Firestone C. The dark room problem. *Trends Cogn Sci (Regul. Ed.)*. 2020;24(5):346–348. doi: [10.1016/j.tics.2020.02.006](https://doi.org/10.1016/j.tics.2020.02.006)
- [61] Chaplot DS, Gandhi D, Gupta S, et al. Learning to explore using active neural slam. In: International Conference on Learning Representations (ICLR). Addis Ababa; 2020.
- [62] Ramakrishnan SK, Jayaraman D, Grauman K. Emergence of exploratory look-around behaviors through active observation completion. *Sci Rob*. 2019;4(30):eaaw6326. doi: [10.1126/scirobotics.aaw6326](https://doi.org/10.1126/scirobotics.aaw6326)
- [63] Ammirato P, Poirson P, Park E, et al. A dataset for developing and benchmarking active vision. In: IEEE International Conference on Robotics and Automation (ICRA). Marina Bay Sands, Singapore; 2017. p. 1378–1385.
- [64] Ognibene D, Demiris Y. Towards active event recognition. In: IJCAI. Beijing; 2013. p. 2495–2501.
- [65] Lee K, Ognibene D, Chang HJ, et al. Stare: spatio-temporal attention relocation for multiple structured activities detection. *IEEE Trans Image Process*. 2015; 24(12):5916–5927. doi: [10.1109/TIP.2015.2487837](https://doi.org/10.1109/TIP.2015.2487837)
- [66] Donnarumma F, Costantini M, Ambrosini E, et al. Action perception as hypothesis testing. *Cortex*. 2017; 89:45–60. doi: [10.1016/j.cortex.2017.01.016](https://doi.org/10.1016/j.cortex.2017.01.016)
- [67] Wang B, Huang L, Hoai M. Active vision for early recognition of human actions. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA; 2020.
- [68] Schwartenbeck P, Passecker J, Hauser TU, et al. Computational mechanisms of curiosity and goal-directed exploration. *Elife*. 2019;8:e41703. doi: [10.7554/eLife.41703](https://doi.org/10.7554/eLife.41703)
- [69] Denzler J, Brown C. Information theoretic sensor data selection for active object recognition and state estimation. *IEEE Trans Pattern Anal Mach Intell*. 2002;24(2):145–157. doi: [10.1109/34.982896](https://doi.org/10.1109/34.982896)
- [70] Sommerlade E, Reid I. Information-theoretic active scene exploration. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Anchorage, AK; 2008. p. 1–7.

- [71] Hafner D, Ortega PA, Ba J, et al. Action and perception as divergence minimization. preprint [arXiv:200901791](https://arxiv.org/abs/200901791). 2020.
- [72] Parr T, Friston KJ. Generalised free energy and active inference. *Biol Cybern.* 2019;113(5):495–513. doi: [10.1007/s00422-019-00805-w](https://doi.org/10.1007/s00422-019-00805-w)
- [73] Attias H. Planning by probabilistic inference. In: International Workshop on Artificial Intelligence and Statistics. Key West, FL: MLR Press; 2003. PMLR R4. p. 9–16.
- [74] Toussaint M. Robot trajectory optimization using approximate inference. In: International Conference on Machine Learning (ICML). Montreal; 2009. p. 1049–1056.
- [75] Kappen HJ, Gómez V, Opper M. Optimal control as a graphical model inference problem. *Mach Learn.* 2012;87(2):159–182. doi: [10.1007/s10994-012-5278-7](https://doi.org/10.1007/s10994-012-5278-7)
- [76] Botvinick M, Toussaint M. Planning as inference. *Trends Cogn Sci (Regul. Ed.)*. 2012;16(10):485–488. doi: [10.1016/j.tics.2012.08.006](https://doi.org/10.1016/j.tics.2012.08.006)
- [77] Millidge B, Tschantz A, Seth AK, et al. On the relationship between active inference and control as inference. In: International Workshop on Active Inference. Ghent: Springer; 2020. p. 3–11.
- [78] Van de Cruys S, Friston K, Clark A. Controlled optimism: reply to sun and firestone on the dark room problem. *Trends Cogn Sci (Regul. Ed.)*. 2020;24(9):1–2. doi: [10.1016/j.tics.2020.05.012](https://doi.org/10.1016/j.tics.2020.05.012)
- [79] Levine S. Reinforcement learning and control as probabilistic inference: tutorial and review. preprint [arXiv:1805.00909](https://arxiv.org/abs/1805.00909). 2018.
- [80] Haarnoja T, Zhou A, Abbeel P, et al. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: International Conference on Machine Learning (ICML). Stockholm; 2018. p. 1861–1870.
- [81] Okada M, Taniguchi T. Dreaming: model-based reinforcement learning by latent imagination without reconstruction. In: IEEE International Conference on Robotics and Automation (ICRA). Xi'an; 2021. p. 4209–4215.
- [82] Okumura R, Nishio N, Taniguchi T. Tactile-sensitive NewtonianVAE for high-accuracy industrial connector-socket insertion. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Kyoto; 2022.
- [83] Kaiser L, Babaeizadeh M, Milos P, et al. Model-based reinforcement learning for Atari. preprint [arXiv:1903.0374](https://arxiv.org/abs/1903.0374). 2019.
- [84] Ke NR, Singh A, Touati A, et al. Learning dynamics model in reinforcement learning by incorporating the long term future. preprint [arXiv:1903.01599](https://arxiv.org/abs/1903.01599). 2019.
- [85] Okada M, Kosaka N, Taniguchi T. Planet of the Bayesians: reconsidering and improving deep planning network by incorporating bayesian inference. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Virtual; 2020. p. 5611–5618.
- [86] Okada M, Taniguchi T. DreamingV2: reinforcement learning with discrete world models without reconstruction. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Kyoto; 2022.
- [87] Jaques M, Burke M, Hospedales TM. NewtonianVAE: proportional control and goal identification from pixels via physical latent spaces. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN; 2021. p. 4454–4463.
- [88] Tani J. Learning to generate articulated behavior through the bottom-up and the top-down interaction processes. *Neural Netw.* 2003;16(1):11–23. doi: [10.1016/S0893-6080\(02\)00214-9](https://doi.org/10.1016/S0893-6080(02)00214-9)
- [89] Yamashita Y, Tani J. Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment. *PLoS Comput Biol.* 2008;4(11):e1000220. doi: [10.1371/journal.pcbi.1000220](https://doi.org/10.1371/journal.pcbi.1000220)
- [90] Nishimoto R, Tani J. Development of hierarchical structures for actions and motor imagery: a constructivist view from synthetic neuro-robotics study. *Psychol Res.* 2009;73(4):545–558. doi: [10.1007/s00426-009-0236-0](https://doi.org/10.1007/s00426-009-0236-0)
- [91] Namikawa J, Nishimoto R, Tani J. A neurodynamic account of spontaneous behaviour. *PLoS Comput Biol.* 2011;7(10):e1002221. doi: [10.1371/journal.pcbi.1002221](https://doi.org/10.1371/journal.pcbi.1002221)
- [92] Yamashita Y, Tani J. Spontaneous prediction error generation in schizophrenia. *PLoS ONE.* 2012;7(5):e37843. doi: [10.1371/journal.pone.0037843](https://doi.org/10.1371/journal.pone.0037843)
- [93] Murata S, Namikawa J, Arie H, et al. Learning to reproduce fluctuating time series by inferring their time-dependent stochastic properties: application in robot learning via tutoring. *IEEE Trans Auton Ment Dev.* 2013;5(4):298–310. doi: [10.1109/TAMD.4563672](https://doi.org/10.1109/TAMD.4563672)
- [94] Murata S, Yamashita Y, Arie H, et al. Learning to perceive the world as probabilistic or deterministic via interaction with others: a neuro-robotics experiment. *IEEE Trans Neural Netw Learn Syst.* 2017;28(4):830–848. doi: [10.1109/TNNLS.2015.2492140](https://doi.org/10.1109/TNNLS.2015.2492140)
- [95] Ahmadi A, Tani J. A novel predictive-coding-inspired variational RNN model for online prediction and recognition. *Neural Comput.* 2019;31(11):2025–2074. doi: [10.1162/neco_a_01228](https://doi.org/10.1162/neco_a_01228)
- [96] Lanillos P, Oliva D, Philippsen A, et al. A review on neural network models of schizophrenia and autism spectrum disorder. *Neural Netw.* 2020;122:338–363. doi: [10.1016/j.neunet.2019.10.014](https://doi.org/10.1016/j.neunet.2019.10.014)
- [97] Van de Cruys S, Evers K, Van der Hallen R, et al. Precise minds in uncertain worlds: predictive coding in autism. *Psychol Rev.* 2014;121(4):649–675. doi: [10.1037/a0037665](https://doi.org/10.1037/a0037665)
- [98] Lawson RP, Rees G, Friston KJ. An aberrant precision account of autism. *Front Hum Neurosci.* 2014;8:1–10. doi: [10.3389/fnhum.2014.00302](https://doi.org/10.3389/fnhum.2014.00302)
- [99] Idei H, Murata S, Chen Y, et al. A neurorobotics simulation of autistic behavior induced by unusual sensory precision. *Comput Psychiatry.* 2018;2:164–182. doi: [10.1162/CPSY_a_00019](https://doi.org/10.1162/CPSY_a_00019)
- [100] Idei H, Murata S, Yamashita Y, et al. Homogeneous intrinsic neuronal excitability induces overfitting to sensory noise: a robot model of neurodevelopmental disorder. *Front Psychiatry.* 2020;11:1–15. doi: [10.3389/fpsy.2020.00762](https://doi.org/10.3389/fpsy.2020.00762)
- [101] Idei H, Murata S, Yamashita Y, et al. Paradoxical sensory reactivity induced by functional disconnection in a robot model of neurodevelopmental disorder. *Neural Netw.* 2021;138:150–163. doi: [10.1016/j.neunet.2021.01.033](https://doi.org/10.1016/j.neunet.2021.01.033)
- [102] Sancaktar C, van Gerven M, Lanillos P. End-to-end pixel-based deep active inference for body perception and action. In: Joint IEEE International Conference on

- Development and Learning and Epigenetic Robotics (ICDL-EpiRob). Online; 2020.
- [103] Pezzato C, Baioumy M, Corbato CH, et al. Active inference for fault tolerant control of robot manipulators with sensory faults. In: International workshop on active inference. Ghent: Springer; 2020. p. 20–27.
- [104] Meo C, Lanillos P. Multimodal vae active inference controller. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Prague: IEEE; 2021.
- [105] Lanillos P, Cheng G. Active inference with function learning for robot body perception. In: International Workshop on Continual Unsupervised Sensorimotor Learning, IEEE Developmental Learning and Epigenetic Robotics (ICDL-EpiRob). Tokyo; 2018.
- [106] Lanillos P, Cheng G. Adaptive robot body learning and estimation through predictive coding. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Madrid: IEEE; 2018. p. 4083–4090.
- [107] Ito M, Noda K, Hoshino Y, et al. Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model. *Neural Netw.* 2006;19(3):323–337. doi: [10.1016/j.neunet.2006.02.007](https://doi.org/10.1016/j.neunet.2006.02.007)
- [108] Sugita Y, Tani J. Learning semantic combinatoriality from the interaction between linguistic and behavioral processes. *Adapt Behav.* 2005;13(1):33–52. doi: [10.1177/105971230501300102](https://doi.org/10.1177/105971230501300102)
- [109] Chen Y, Murata S, Arie H, et al. Emergence of interactive behaviors between two robots by prediction error minimization mechanism. In: Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob). Cergy-Pontoise: IEEE; 2016. p. 302–307.
- [110] Çatal O, Verbelen T, Van de Maele T, et al. Robot navigation as hierarchical active inference. *Neural Netw.* 2021;142:192–204. doi: [10.1016/j.neunet.2021.05.010](https://doi.org/10.1016/j.neunet.2021.05.010)
- [111] Imre M, Oztop E, Nagai Y, et al. Affordance-based altruistic robotic architecture for human–robot collaboration. *Adapt Behav.* 2019;27(4):223–241. doi: [10.1177/1059712318824697](https://doi.org/10.1177/1059712318824697)
- [112] Seker MY, Ahmetoglu A, Nagai Y, et al. Imitation and mirror systems in robots through deep modality blending networks. *Neural Netw.* 2022;146:22–35. doi: [10.1016/j.neunet.2021.11.004](https://doi.org/10.1016/j.neunet.2021.11.004)
- [113] Seker MY, Imre M, Piater J, et al. Conditional neural movement primitives. In: Robotics Science and Systems (RSS). Messe Freiburg; 2019.
- [114] Tschantz A, Baltieri M, Seth AK, et al. Scaling active inference. In: International Joint Conference on Neural Networks (IJCNN). Glasgow: IEEE; 2020. p. 1–8.
- [115] van Bergen RS, Lanillos P. Object-based active inference. In: Buckley CL, Cialfi D, Lanillos P, Ramstead M, Sajid N, Shimazaki H, Verbelen T, editors. Active inference. Cham: Springer Nature Switzerland; 2023. p. 50–64.
- [116] Da Costa L, Lanillos P, Sajid N, et al. How active inference could help revolutionise robotics. *Entropy.* 2022;24(3):361. doi: [10.3390/e24030361](https://doi.org/10.3390/e24030361)
- [117] Valenzo D, Ciria A, Schillaci G, et al. Grounding context in embodied cognitive robotics. *Front Neurobot.* 2022;16. doi: [10.3389/fnbot.2022.843108](https://doi.org/10.3389/fnbot.2022.843108)
- [118] Fountas Z, Sajid N, Mediano P, et al. Deep active inference agents using monte-carlo methods. *Adv Neural Inf Process Syst.* 2020;33:11662–11675.
- [119] Mazzaglia P, Verbelen T, Dhoedt B. Contrastive active inference. *Adv Neural Inf Process Syst.* 2021;34:13870–13882.
- [120] Böhmer W, Springenberg JT, Boedecker J, et al. Autonomous learning of state representations for control: an emerging field aims to autonomously learn state representations for reinforcement learning agents from their real-world sensor observations. *KI-Künstliche Intelligenz.* 2015;29(4):353–362. doi: [10.1007/s13218-015-0356-1](https://doi.org/10.1007/s13218-015-0356-1)
- [121] Achille A, Soatto S. A separation principle for control in the age of deep learning. *Annu Rev Control Robotics Autonomous Syst.* 2018;1:287–307. doi: [10.1146/control.2018.1.issue-1](https://doi.org/10.1146/control.2018.1.issue-1)
- [122] Lesort T, Díaz-Rodríguez N, Goudou JF, et al. State representation learning for control: an overview. *Neural Netw.* 2018;108:379–392. doi: [10.1016/j.neunet.2018.07.006](https://doi.org/10.1016/j.neunet.2018.07.006)
- [123] Stoianov I, Maisto D, Pezzulo G. The hippocampal formation as a hierarchical generative model supporting generative replay and continual learning. *Prog Neurobiol.* 2022;217:102329. doi: [10.1016/j.pneurobio.2022.102329](https://doi.org/10.1016/j.pneurobio.2022.102329)
- [124] Gumbsch C, Butz MV, Martius G. Sparsely changing latent states for prediction and planning in partially observable domains. *Adv Neural Inf Process Syst.* 2021;34:17518–17531.
- [125] Greff K, Kaufman RL, Kabra R, et al. Multi-object representation learning with iterative variational inference. In: International Conference on Machine Learning (ICML). Long Beach, CA; 2019. p. 2424–2433.
- [126] Wang T, Du SS, Torralba A, et al. Denoised MDPs: learning world models better than the world itself. preprint [arXiv:220615477](https://arxiv.org/abs/220615477). 2022.
- [127] Watter M, Springenberg J, Boedecker J, et al. Embed to control: a locally linear latent dynamics model for control from raw images. *Adv Neural Inf Process Syst.* 2015;28:1–9.
- [128] Levine S, Koltun V. Guided policy search. In: International Conference on Machine Learning. Atlanta: PMLR; 2013;28(3):1–9.
- [129] Hansen N. The CMA evolution strategy: a tutorial. preprint [arXiv:160400772](https://arxiv.org/abs/160400772). 2016.
- [130] Rubinstein RY. Optimization of computer simulation models with rare events. *Eur J Oper Res.* 1997;99(1):89–112. doi: [10.1016/S0377-2217\(96\)00385-2](https://doi.org/10.1016/S0377-2217(96)00385-2)
- [131] Zhang M, Vikram S, Smith L, et al. Solar: deep structured representations for model-based reinforcement learning. In: International Conference on Machine Learning. PMLR; 2019. p. 7444–7453.
- [132] Racanière S, Weber T, Reichert D, et al. Imagination-augmented agents for deep reinforcement learning. *Adv Neural Inf Process Syst.* 2017;30:1–12.
- [133] Wang T, Bao X, Clavera I, et al. Benchmarking model-based reinforcement learning. preprint [arXiv:190702057](https://arxiv.org/abs/190702057). 2019.
- [134] Paster K, McIlraith SA, Ba J. Planning from pixels using inverse dynamics models. preprint [arXiv:201202419](https://arxiv.org/abs/201202419). 2020.

- [135] Hafner D, Pasukonis J, Ba J, et al. Mastering diverse domains through world models. preprint [arXiv:230104104](https://arxiv.org/abs/230104104). 2023.
- [136] Köhler W, Winter E. The mentality of apes: international library of psychology, philosophy, and scientific method. London: Routledge & Kegan Paul Ltd; 1925.
- [137] Steedman M. Plans, affordances, and combinatory grammar. *Linguist Philos.* 2002;25(5-6):723–753. doi: [10.1023/A:1020820000972](https://doi.org/10.1023/A:1020820000972)
- [138] Goel V, Gold B, Kapur S, et al. Neuroanatomical correlates of human reasoning. *J Cogn Neurosci.* 1998;10(3):293–302. doi: [10.1162/089892998562744](https://doi.org/10.1162/089892998562744)
- [139] Al N, Nolen-Hoeksema S. Atkinson and Hilgard's introduction to psychology. Andover: Cengage Learning EMEA; 2014.
- [140] Russell SJ, Norvig P. Artificial intelligence: a modern approach. London: Pearson Education Limited; 2016.
- [141] Mugan J, Kuipers B. Autonomous learning of high-level states and actions in continuous environments. *IEEE Trans Auton Ment Dev.* 2012;4(1):70–86. doi: [10.1109/TAMD.2011.2160943](https://doi.org/10.1109/TAMD.2011.2160943)
- [142] Ahmetoglu A, Ugur E, Asada M, et al. High-level features for resource economy and fast learning in skill transfer. *Adv Robot.* 2022;36(5-6):291–303. doi: [10.1080/01691864.2021.2019613](https://doi.org/10.1080/01691864.2021.2019613)
- [143] Konidaris G, Kaelbling LP, Lozano-Perez T. Constructing symbolic representations for high-level planning. In: *AAAI Conference on Artificial Intelligence (AAAI)*. Québec City; 2014: 28(1).
- [144] Konidaris G, Kaelbling L, Lozano-Perez T. Symbol acquisition for probabilistic high-level planning. In: *International Joint Conference on Artificial Intelligence (IJCAI)*. Buenos Aires: AAAI Press; 2015.
- [145] James S, Rosman B, Konidaris G. Learning portable representations for high-level planning. preprint [arXiv:190512006](https://arxiv.org/abs/190512006). 2019.
- [146] Ugur E, Piater J. Bottom-up learning of object categories, action effects and logical rules: From continuous manipulative exploration to symbolic planning. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Seattle, WA: IEEE; 2015. p. 2627–2633.
- [147] Ugur E, Piater J. Refining discovered symbols with multi-step interaction experience. In: *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. Seoul: IEEE; 2015. p. 1007–1012.
- [148] Ahmetoglu A, Seker MY, Piater J, et al. DeepSym: deep symbol generation and rule learning from unsupervised continuous robot interaction for planning. *J Artificial Intell Res.* 2022;75:709–745. doi: [10.1613/jair.1.13754](https://doi.org/10.1613/jair.1.13754)
- [149] Şahin E, Cakmak M, Doğar MR, et al. To afford or not to afford: a new formalization of affordances toward affordance-based robot control. *Adapt Behav.* 2007;15(4):447–472. doi: [10.1177/1059712307084689](https://doi.org/10.1177/1059712307084689)
- [150] Ahmetoglu A, Oztop E, Ugur E. Learning multi-object symbols for manipulation with attentive deep effect predictors. preprint [arXiv:220801021](https://arxiv.org/abs/220801021). 2022.
- [151] Asai M, Fukunaga A. Classical planning in deep latent space: bridging the subsymbolic-symbolic boundary. preprint [arXiv:170500154](https://arxiv.org/abs/170500154). 2017.
- [152] Asai M, Muise C. Learning neural-symbolic descriptive planning models via cube-space priors: the voyage home (to strips). preprint [arXiv:200412850](https://arxiv.org/abs/200412850). 2020.
- [153] Gibson JJ. The ecological approach to visual perception. Boston: Houghton Mifflin; 1979.
- [154] Zech P, Haller S, Lakani SR, et al. Computational models of affordance in robotics: a taxonomy and systematic classification. *Adapt Behav.* 2017;25(5):235–271. doi: [10.1177/1059712317726357](https://doi.org/10.1177/1059712317726357)
- [155] Jamone L, Ugur E, Cangelosi A, et al. Affordances in psychology, neuroscience and robotics: a survey. *IEEE Trans Cognitive Develop Syst.* 2016;10(1):4–25. doi: [10.1109/TCDS.2016.2594134](https://doi.org/10.1109/TCDS.2016.2594134)
- [156] Renaudo E, Zech P, Chatila R, et al. Computational models of affordance for robotics. *Front Neurobot.* 2022;16:1045355. doi: [10.3389/fnbot.2022.1045355](https://doi.org/10.3389/fnbot.2022.1045355)
- [157] Gibson JJ. The senses considered as perceptual systems. Boston: Houghton Mifflin; 1966.
- [158] Gibson JJ. The theory of affordances. In: *Perceiving, acting, and knowing: toward an ecological psychology*. Minnesota: Lawrence Erlbaum Associates; 1977.
- [159] Norman DA. Affordance, conventions, and design. *Interactions.* 1999;6(3):38–42. doi: [10.1145/301153.301168](https://doi.org/10.1145/301153.301168)
- [160] Gibson EJ. An odyssey in learning and perception. Cambridge, MA: MIT Press; 1994.
- [161] Gibson EJ. Perceptual learning in development: some basic concepts. *Ecol Psychol.* 2000;12(4):295–302. doi: [10.1207/S15326969ECO1204_04](https://doi.org/10.1207/S15326969ECO1204_04)
- [162] Gibson EJ. The world is so full of a number of things: on specification and perceptual learning. *Ecol Psychol.* 2003;15(4):283–288. doi: [10.1207/s15326969eco1504_3](https://doi.org/10.1207/s15326969eco1504_3)
- [163] Montesano L, Lopes M, Bernardino A, et al. Learning object affordances: from sensory-motor coordination to imitation. *IEEE Trans Robot.* 2008;24(1):15–26. doi: [10.1109/TRO.2007.914848](https://doi.org/10.1109/TRO.2007.914848)
- [164] Ugur E, Şahin E, Oztop E. Unsupervised learning of object affordances for planning in a mobile manipulation platform. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Shanghai: IEEE; 2011. p. 4312–4317.
- [165] Krueger N, Geib C, Piater J, et al. Object-action complexes: grounded abstractions of sensory-motor processes. *Rob Auton Syst.* 2011;59(10):740–757. doi: [10.1016/j.robot.2011.05.009](https://doi.org/10.1016/j.robot.2011.05.009)
- [166] Szedmak S, Ugur E, Piater J. Knowledge propagation and relation learning for predicting action effects. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Chicago, IL: IEEE; 2014. p. 623–629.
- [167] Gonçalves A, Abrantes J, Saponaro G, et al. Learning intermediate object affordances: towards the development of a tool concept. In: *4th International Conference on Development and Learning and on Epigenetic Robotics*. Genoa: IEEE; 2014. p. 482–488.
- [168] Dehban A, Jamone L, Kampff AR, et al. Denoising auto-encoders for learning of objects and tools affordances in continuous space. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Stockholm: IEEE; 2016. p. 4866–4871.
- [169] Dehban A, Jamone L, Kampff AR, et al. A deep probabilistic framework for heterogeneous self-supervised learning of affordances. In: *IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)*. Birmingham: IEEE; 2017. p. 476–483.

- [170] Ugur E, Piater J. Emergent structuring of interdependent affordance learning tasks using intrinsic motivation and empirical feature selection. *IEEE Trans Cognitive Develop Syst.* 2017;9(4):328–340. doi: [10.1109/TCDS.2016.2581307](https://doi.org/10.1109/TCDS.2016.2581307)
- [171] Sarathy V, Scheutz M. A logic-based computational framework for inferring cognitive affordances. *IEEE Trans Cognitive Develop Syst.* 2018;10(1):26–43. doi: [10.1109/TCDS.2016.2615326](https://doi.org/10.1109/TCDS.2016.2615326)
- [172] Stramandinoli F, Tikhanoff V, Pattacini U, et al. Heteroscedastic regression and active learning for modeling affordances in humanoids. *IEEE Trans Cognitive Develop Syst.* 2018;10(2):455–468. doi: [10.1109/TCDS.2017.2700207](https://doi.org/10.1109/TCDS.2017.2700207)
- [173] Tekden AE, Erdem A, Erdem E, et al. Belief regulated dual propagation nets for learning action effects on groups of articulated objects. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Paris: IEEE; 2020. p. 10556–10562.
- [174] Antunes A, Jamone L, Saponaro G, et al. From human instructions to robot actions: formulation of goals, affordances and probabilistic planning. In: *IEEE ICRA*; Stockholm: IEEE; 2016.
- [175] Dehban A, Zhang S, Cauli N, et al. Learning deep features for robotic inference from physical interactions. *IEEE Trans Cognitive Develop Syst.* 2022. doi: [10.1109/TCDS.2022.3152383](https://doi.org/10.1109/TCDS.2022.3152383)
- [176] Shen Z, Wu Y. Investigation of practical use of humanoid robots in elderly care centres. In: *Proceedings of the 4th International Conference on Human Agent Interaction*. Biopolis: Association for Computing Machinery New York, USA; 2016. p. 63–66.
- [177] Albrecht SV, Stone P. Autonomous agents modelling other agents: a comprehensive survey and open problems. *Artif Intell.* 2018;258:66–95. doi: [10.1016/j.artint.2018.01.002](https://doi.org/10.1016/j.artint.2018.01.002)
- [178] El Zaatari S, Marei M, Li W, et al. Cobot programming for collaborative industrial tasks: an overview. *Rob Auton Syst.* 2019;116:162–180. doi: [10.1016/j.robot.2019.03.003](https://doi.org/10.1016/j.robot.2019.03.003)
- [179] Hentout A, Aouache M, Maoudj A, et al. Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017. *Adv Robot.* 2019;33(15-16):764–799. doi: [10.1080/01691864.2019.1636714](https://doi.org/10.1080/01691864.2019.1636714)
- [180] Magrini E, Ferraguti F, Ronga AJ, et al. Human-robot coexistence and interaction in open industrial cells. *Robot Comput Integr Manuf.* 2020;61:101846. doi: [10.1016/j.rcim.2019.101846](https://doi.org/10.1016/j.rcim.2019.101846)
- [181] Ognibene D, Foulsham T, Marchegiani L, et al. Active vision and perception in human-robot collaboration. *Front Neurorobot.* 2022;16. doi: [10.3389/fnbot.2022.848065](https://doi.org/10.3389/fnbot.2022.848065)
- [182] Semeraro F, Griffiths A, Cangelosi A. Human–robot collaboration and machine learning: A systematic review of recent research. *Robot Comput Integr Manuf.* 2023;79:102432. doi: [10.1016/j.rcim.2022.102432](https://doi.org/10.1016/j.rcim.2022.102432)
- [183] Ognibene D, Chinellato E, Sarabia M, et al. Contextual action recognition and target localization with an active allocation of attention on a humanoid robot. *Bioinspir Biomim.* 2013;8(3):035002. doi: [10.1088/1748-3182/8/3/035002](https://doi.org/10.1088/1748-3182/8/3/035002)
- [184] Fotopoulou A, Tsakiris M. Mentalizing homeostasis: the social origins of interoceptive inference. *Neuropsychanalysis.* 2017;19(1):3–28. doi: [10.1080/15294145.2017.1294031](https://doi.org/10.1080/15294145.2017.1294031)
- [185] Veissière SP, Constant A, Ramstead MJ, et al. Thinking through other minds: a variational approach to cognition and culture. *Behavioral Brain Sci.* 2020;43:e90. doi: [10.1017/S0140525X190012133](https://doi.org/10.1017/S0140525X190012133)
- [186] Saxe R. Uniquely human social cognition. *Curr Opin Neurobiol.* 2006;16(2):235–239. doi: [10.1016/j.conb.2006.03.001](https://doi.org/10.1016/j.conb.2006.03.001)
- [187] Bianco F, Ognibene D. Functional advantages of an adaptive theory of mind for robotics: a review of current architectures. In: *2019 11th Computer Science and Electronic Engineering (CEECE)*. Computer Sci Electronic Eng (CEECE). Colchester: IEEE; 2019. doi: [10.1109/CEECE47804.2019.8974334](https://doi.org/10.1109/CEECE47804.2019.8974334)
- [188] Baker CL, Jara-Ettinger J, Saxe R, et al. Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour.* 2017;1(4):1–10. doi: [10.1038/s41562-017-0064](https://doi.org/10.1038/s41562-017-0064)
- [189] Bianco F, Ognibene D. Robot learning theory of mind through self-observation: exploiting the intentions-beliefs synergy. preprint [arXiv:221009435](https://arxiv.org/abs/221009435). 2022.
- [190] Ognibene D, Mirante L, Marchegiani L. Proactive intention recognition for joint human-robot search and rescue missions through monte-carlo planning in POMDP environments. In: *International Conference on Social Robotics*. Madrid: Springer; 2019. p. 332–343.
- [191] Bianchi F, Marelli M, Nicoli P, et al. Sweat: scoring polarization of topics across different corpora. preprint [arXiv:210907231](https://arxiv.org/abs/210907231). 2021.
- [192] Roggen D, Calatroni A, Rossi M, et al. Collecting complex activity datasets in highly rich networked sensor environments. In: *International Conference on Networked Sensing Systems (INSS)*. Kassel: IEEE; 2010. p. 233–240.
- [193] Zeng M, Nguyen LT, Yu B, et al. Convolutional neural networks for human activity recognition using mobile sensors. In: *6th International Conference on Mobile Computing, Applications and Services*. Kassel: IEEE; 2014. p. 197–205.
- [194] Yang J, Nguyen MN, San PP, et al. Deep convolutional neural networks on multichannel time series for human activity recognition. In: *24th International Joint Conference on Artificial Intelligence*. Buenos Aires: AAAI Press; 2015. p. 3995–4001.
- [195] Lee SU, Hofmann A, Williams B. A model-based human activity recognition for human–robot collaboration. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Macau: IEEE; 2019. p. 736–743.
- [196] Ramírez M, Geffner H. Plan recognition as planning. In: *International Joint Conference on Artificial Intelligence (IJCAI)*. Pasadena, CA: Morgan Kaufmann Publishers Inc; 2009. p. 1778–1783.
- [197] Baker CL, Saxe R, Tenenbaum JB. Action understanding as inverse planning. *Cognition.* 2009;113(3):329–349. doi: [10.1016/j.cognition.2009.07.005](https://doi.org/10.1016/j.cognition.2009.07.005)
- [198] Sohrabi S, Riabov AV, Udrea O. Plan recognition as planning revisited. In: *International Joint Conference*

- on Artificial Intelligence (IJCAI). New York, NY: AAAI Press; 2016. p. 3258–3264.
- [199] Ramirez M, Geffner H. Goal recognition over POMDPs: Inferring the intention of a pomdp agent. In: International Joint Conference on Artificial Intelligence (IJCAI) Vol. 3. Barcelona: AAAI Press; 2011. p. 2009–2014.
- [200] Baker C, Saxe R, Tenenbaum J. Bayesian theory of mind: modeling joint belief-desire attribution. In: Proceedings of the Annual Meeting of The Cognitive Science Society. Vol. 33. Boston, MA; 2011. p. 2469–2474.
- [201] Shvo M, McIlraith SA. Active goal recognition. In: Proceedings of the AAAI Conference on Artificial Intelligence. New York, NY: AAAI Press; 2020. Vol. 34. p. 9957–9966.
- [202] Amato C, Baisero A. Active goal recognition. preprint [arXiv:190911173](https://arxiv.org/abs/190911173). 2019.
- [203] Demiris Y, Simmons G. Perceiving the unusual: temporal properties of hierarchical motor representations for action perception. *Neural Netw.* 2006;19(3):272–284. doi: [10.1016/j.neunet.2006.02.005](https://doi.org/10.1016/j.neunet.2006.02.005)
- [204] Cardona-Rivera RE, Young RM. Toward combining domain theory and recipes in plan recognition. In: Workshops at the 31 AAAI Conference on Artificial Intelligence. San Francisco, CA: AAAI Press; 2017.
- [205] Proietti R, Pezzulo G, Tessari A. An active inference model of hierarchical action understanding, learning and imitation. *Physics of Life Reviews.* 2023. doi: [10.1016/j.plrev.2023.05.012](https://doi.org/10.1016/j.plrev.2023.05.012)
- [206] Griffiths TL, Lieder F, Goodman ND. Rational use of cognitive resources: levels of analysis between the computational and the algorithmic. *Top Cogn Sci.* 2015;7(2):217–229. doi: [10.1111/tops.12142](https://doi.org/10.1111/tops.12142)
- [207] Ognibene D, Fiore VG, Gu X. Addiction beyond pharmacological effects: the role of environment complexity and bounded rationality. *Neural Netw.* 2019; 116:269–278. doi: [10.1016/j.neunet.2019.04.022](https://doi.org/10.1016/j.neunet.2019.04.022)
- [208] Ziebart BD, Maas AL, Bagnell JA, et al. Maximum entropy inverse reinforcement learning. In: AAAI. Vol. 8. Chicago, IL, USA: AAAI Press; 2008. p. 1433–1438.
- [209] Zhi-Xuan T, Mann J, Silver T, et al. Online bayesian goal inference for boundedly rational planning agents. In: Larochelle H, Ranzato M, Hadsell R, Balcan M, Lin H, editors. Advances in neural information processing systems. Vol. 33. Virtual-only Conference, Online: Curran Associates, Inc.; 2020. p. 19238–19250. <https://proceedings.neurips.cc/paper/2020/file/df3aebc649f9e3b674eeb790a4da224e-Paper.pdf>.
- [210] Johnson M, Demiris Y. Perceptual perspective taking and action recognition. *Int J Advanced Rob Syst.* 2005;2(4):32. doi: [10.5772/5775](https://doi.org/10.5772/5775)
- [211] Pandey AK, Ali M, Alami R. Towards a task-aware proactive sociable robot based on multi-state perspective-taking. *Int J Soc Robot.* 2013;5(2):215–236. doi: [10.1007/s12369-013-0181-3](https://doi.org/10.1007/s12369-013-0181-3)
- [212] Fischer T, Demiris Y. Computational modeling of embodied visual perspective taking. *IEEE Trans Cognitive Develop Syst.* 2019;12(4):723–732. doi: [10.1109/TCDS](https://doi.org/10.1109/TCDS)
- [213] Torabi F, Warnell G, Stone P. Recent advances in imitation learning from observation. *arXiv e-prints. arXiv-1905*. 2019.
- [214] Mörtl A, Lawitzky M, Kucukyilmaz A, et al. The role of roles: physical cooperation between humans and robots. *Int J Rob Res.* 2012;31(13):1656–1674. doi: [10.1177/0278364912455366](https://doi.org/10.1177/0278364912455366)
- [215] Li Y, Tee KP, Yan R, et al. A framework of human–robot coordination based on game theory and policy iteration. *IEEE Trans Robot.* 2016;32(6):1408–1418. doi: [10.1109/TRO.2016.2597322](https://doi.org/10.1109/TRO.2016.2597322)
- [216] Pezzulo G, Donnarumma F, Dindo H, et al. The body talks: sensorimotor communication and its brain and kinematic signatures. *Phys Life Rev.* 2019;28:1–21. doi: [10.1016/j.plrev.2018.06.014](https://doi.org/10.1016/j.plrev.2018.06.014)
- [217] Ognibene D, Giglia G, Marchegiani L, et al. Implicit perception simplicity and explicit perception complexity in sensorimotor communication. *Phys Life Rev.* 2019;28:36–38. doi: [10.1016/j.plrev.2019.01.017](https://doi.org/10.1016/j.plrev.2019.01.017)
- [218] Kok BC, Soh H. Trust in robots: challenges and opportunities. *Current Robot Rep.* 2020;1(4):297–309. doi: [10.1007/s43154-020-00029-y](https://doi.org/10.1007/s43154-020-00029-y)
- [219] Craik K. The nature of explanation. Cambridge: Cambridge University Press; 1943.
- [220] von Helmholtz H. *Handbuch der physiologischen optik*. Leipzig: L. Voss; 1867.
- [221] Gregory RL. Knowledge in perception and illusion. *Philos Trans Royal Soc B: Biological Sci.* 1997;352: 1121–1128. doi: [10.1098/rstb.1997.0095](https://doi.org/10.1098/rstb.1997.0095)
- [222] Doya K, Ishii S, Pouget A, et al. Bayesian brain: probabilistic approaches to neural coding. 1st ed. Cambridge, MA: The MIT Press; 2007.
- [223] Shiffrin RM, Bassett DS, Kriegeskorte N, et al. The brain produces mind by modeling. *Proc Natl Acad Sci.* 2020;117(47):29299–29301. doi: [10.1073/pnas.1912340117](https://doi.org/10.1073/pnas.1912340117)
- [224] Friston K, FitzGerald T, Rigoli F, et al. Active inference: a process theory. *Neural Comput.* 2017;29(1):1–49. doi: [10.1162/NECO_a_00912](https://doi.org/10.1162/NECO_a_00912)
- [225] Ma WJ, Beck JM, Latham PE, et al. Bayesian inference with probabilistic population codes. *Nat Neurosci.* 2006;9(11):1432–1438. doi: [10.1038/nn1790](https://doi.org/10.1038/nn1790)
- [226] Buesing L, Bill J, Nessler B, et al. Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS Comput Biol.* 2011;7(11):e1002211. doi: [10.1371/journal.pcbi.1002211](https://doi.org/10.1371/journal.pcbi.1002211)
- [227] Sebastian S, Seemiller ES, Geisler WS. Local reliability weighting explains identification of partially masked objects in natural images. *Proc Natl Acad Sci.* 2020;117(47):29363–29370. doi: [10.1073/pnas.1912331117](https://doi.org/10.1073/pnas.1912331117)
- [228] Lynn CW, Bassett DS. How humans learn and represent networks. *Proc Natl Acad Sci.* 2020;117(47):29407–29415. doi: [10.1073/pnas.1912328117](https://doi.org/10.1073/pnas.1912328117)
- [229] Berkes P, Orban G, Lengyel M, et al. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science.* 2011;331(6013): 83–87. doi: [10.1126/science.1195870](https://doi.org/10.1126/science.1195870)
- [230] Ekman M, Kok P, de Lange FP. Time-compressed preplay of anticipated events in human primary visual cortex.

- Nat Commun. 2017;8(1):1–9. doi: [10.1038/ncomms15276](https://doi.org/10.1038/ncomms15276)
- [231] Daw ND, Gershman SJ, Seymour B, et al. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 2011;69(6):1204–1215. doi: [10.1016/j.neuron.2011.02.027](https://doi.org/10.1016/j.neuron.2011.02.027)
- [232] Gómez CM, Arjona A, Donnarumma F, et al. Tracking the time course of bayesian inference with event related potentials: a study using the central cue posner paradigm. *Front Psychol*. 2019;10:1424. doi: [10.3389/fpsyg.2019.01424](https://doi.org/10.3389/fpsyg.2019.01424)
- [233] Muckli L, De Martino F, Vizioli L, et al. Contextual feedback to superficial layers of v1. *Curr Biol*. 2015;25(20):2690–2695. doi: [10.1016/j.cub.2015.08.057](https://doi.org/10.1016/j.cub.2015.08.057)
- [234] Walsh KS, McGovern DP, Clark A, et al. Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Ann NY Acad Sci*. 2020;1464(1):242–268. doi: [10.1111/nyas.v1464.1](https://doi.org/10.1111/nyas.v1464.1)
- [235] Keller GB, Mscis-Flogel TD. Predictive processing: a canonical cortical computation. *Neuron*. 2018;100(2):424–435. doi: [10.1016/j.neuron.2018.10.003](https://doi.org/10.1016/j.neuron.2018.10.003)
- [236] Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. 2009;10(3):186–198. doi: [10.1038/nrn2575](https://doi.org/10.1038/nrn2575)
- [237] Laird JE, Lebiere C, Rosenbloom PS. A standard model of the mind: toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. *AI Magaz*. 2017;38(4):13–26. doi: [10.1609/aimag.v38i4.2744](https://doi.org/10.1609/aimag.v38i4.2744)
- [238] Margulies DS, Ghosh SS, Goulas A, et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc Natl Acad Sci*. 2016;113(44):12574–12579. doi: [10.1073/pnas.1608282113](https://doi.org/10.1073/pnas.1608282113)
- [239] Ungerleider L, Haxby J. 'what' and 'where' in the human brain. *Curr Opin Neurobiol*. 1994;4(2):157–165. doi: [10.1016/0959-4388\(94\)90066-3](https://doi.org/10.1016/0959-4388(94)90066-3)
- [240] Taniguchi T, Yamakawa H, Nagai T, et al. A whole brain probabilistic generative model: toward realizing cognitive architectures for developmental robots. *Neural Netw*. 2022;150:293–312. doi: [10.1016/j.neunet.2022.02.026](https://doi.org/10.1016/j.neunet.2022.02.026)
- [241] Singer W. Recurrent dynamics in the cerebral cortex: integration of sensory evidence with stored knowledge. *Proc Natl Acad Sci*. 2021;118(33):e2101043118. doi: [10.1073/pnas.2101043118](https://doi.org/10.1073/pnas.2101043118)
- [242] Pezzulo G, Zorzi M, Corbetta M. The secret life of predictive brains: what's spontaneous activity for? *Trends Cogn Sci (Regul. Ed.)*. 2021;25:730–743. doi: [10.1016/j.tics.2021.05.007](https://doi.org/10.1016/j.tics.2021.05.007)
- [243] György Buzsáki M. *The brain from inside out*. Oxford: Oxford University Press; 2019.
- [244] Parr T, Holmes E, Friston KJ, et al. Cognitive effort and active inference. *Neuropsychologia*. 2023;184:108562. doi: [10.1016/j.neuropsychologia.2023.108562](https://doi.org/10.1016/j.neuropsychologia.2023.108562)
- [245] Pezzulo G, Kemere C, van der Meer M. Internally generated hippocampal sequences as a vantage point to probe future-oriented cognition. *Ann NY Acad Sci*. 2017;1396:144–165. doi: [10.1111/nyas.2017.1396.issue-1](https://doi.org/10.1111/nyas.2017.1396.issue-1)
- [246] Nazli I, Ferrari A, Huber-Huber C, et al. Statistical learning is not error-driven. *bioRxiv*. 2022.
- [247] Pezzulo G, Cisek P. Navigating the affordance landscape: feedback control as a process model of behavior and cognition. *Trends Cogn Sci (Regul. Ed.)*. 2016;20(6):414–424. doi: [10.1016/j.tics.2016.03.013](https://doi.org/10.1016/j.tics.2016.03.013)
- [248] Cisek P. Resynthesizing behavior through phylogenetic refinement. *Attention, Perception, Psychophysics*. 2019;81(7):2265–2287. doi: [10.3758/s13414-019-01760-1](https://doi.org/10.3758/s13414-019-01760-1)
- [249] Lieto A, Bhatt M, Oltramari A, et al. The role of cognitive architectures in general artificial intelligence. *Cognitive Syst Res*. 2018;48:1–3. doi: [10.1016/j.cogsys.2017.08.003](https://doi.org/10.1016/j.cogsys.2017.08.003)
- [250] Kotseruba I, Tsotsos JK. 40 years of cognitive architectures: core cognitive abilities and practical applications. *Artif Intell Rev*. 2020;53(1):17–94. doi: [10.1007/s10462-018-9646-y](https://doi.org/10.1007/s10462-018-9646-y)
- [251] Anderson JR. *How can the human mind occur in the physical universe?* Oxford: Oxford University Press; 2009.
- [252] Laird JE. Extending the soar cognitive architecture. *Front Artificial Intell Appl*. 2008;171:224–235.
- [253] Anderson JR. Human symbol manipulation within an integrated cognitive architecture. In: *Cognitive Science*. Routledge; 2005. p. 313–341.
- [254] Puigbo JY, Pumarola A, Téllez RA. Controlling a general purpose service robot by means of a cognitive architecture. In: *Ceur Workshop Proceedings*; Vol. 1100. Torino; 2013. p. 45–55.
- [255] Mohan S, Laird JE. Learning to play Mario. *Tech Rep CCA-TR-2009-03*. 2009.
- [256] Rosenbloom PS, Demski A, Ustun V. The sigma cognitive architecture and system: towards functionally elegant grand unification. *J Artificial General Intell*. 2016;7(1):1–103. doi: [10.1515/jagi-2016-0001](https://doi.org/10.1515/jagi-2016-0001)
- [257] Damgaard MR, Pedersen R, Bak T. Toward an idiomatic framework for cognitive robotics. *Patterns*. 2022;3(7):100533. doi: [10.1016/j.patter.2022.100533](https://doi.org/10.1016/j.patter.2022.100533)
- [258] Bingham E, Chen JP, Jankowiak M, et al. Pyro: deep universal probabilistic programming. *J Machine Learn Res*. 2019;20(1):973–978. <https://jmlr.org/papers/v20/18-403.html>
- [259] Paige B, van de Meent JW, Desmaison A, et al. Learning disentangled representations with semi-supervised deep generative models. *Adv Neural Inf Process Syst*. 2017;30:5927–5937. <https://papers.nips.cc/paper/7174-learning-disentangled-representations-with-semi-supervised-deep-generative-models>
- [260] Tran D, Kucukelbir A, Dieng AB, et al. Edward: a library for probabilistic modeling, inference, and criticism. preprint [arXiv:161009787](https://arxiv.org/abs/161009787). 2016.
- [261] Sandini G, Mohan V, Sciutti A, et al. Social cognition for human-robot symbiosis—challenges and building blocks. *Front Neurobot*. 2018;12:34. doi: [10.3389/fnbot.2018.00034](https://doi.org/10.3389/fnbot.2018.00034)
- [262] Sun R. *Anatomy of the mind: exploring psychological mechanisms and processes with the clarion cognitive architecture*. Oxford: Oxford University Press; 2016.
- [263] Kahneman D. A perspective on judgment and choice: mapping bounded rationality. *American Psychol*. 2003;58(9):697. doi: [10.1037/0003-066X.58.9.697](https://doi.org/10.1037/0003-066X.58.9.697)

- [264] Vahrenkamp N, Wächter M, Kröhnert M, et al. The robot software framework armarx. *IT-Inform Technol.* 2015;57(2):99–111. doi: [10.1515/itit-2014-1066](https://doi.org/10.1515/itit-2014-1066)
- [265] Nakamura T, Nagai T, Taniguchi T. Serket: an architecture for connecting stochastic models to realize a large-scale cognitive model. *Front Neurobot.* 2018;12:1–16. doi: [10.3389/fnbot.2018.00025](https://doi.org/10.3389/fnbot.2018.00025)
- [266] Taniguchi T, Nakamura T, Suzuki M, et al. Neuroserket: development of integrative cognitive system through the composition of deep probabilistic generative models. *New Generation Comput.* 2020;38:23–48. doi: [10.1007/s00354-019-00084-w](https://doi.org/10.1007/s00354-019-00084-w)
- [267] Straka Z, Svoboda T, Hoffmann M. Precnet: next-frame video prediction based on predictive coding. *IEEE Trans Neural Netw Learn Syst.* 2023. doi: [10.1109/TNNLS.2023.3240857](https://doi.org/10.1109/TNNLS.2023.3240857)
- [268] Ozturku OB, Ugur E, Oztop E. High-level representations through unconstrained sensorimotor learning. In: *International Conference on Development and Learning (ICDL)*. Chile; 2020. p. 1–6.
- [269] Ugur E, Sahin E, Oztop E. Self-discovery of motor primitives and learning grasp affordances. In: *IEEE International Conference on Intelligent Robots and Systems (IROS)*. Vilamoura-Algarve: IEEE; 2012. p. 3260–3267.
- [270] Devlin J, Chang MW, Lee K, et al. BERT: pre-training of deep bidirectional transformers for language understanding. preprint [arXiv:181004805](https://arxiv.org/abs/1810.04805). 2019.
- [271] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst.* 2020;33:1877–1901. https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf
- [272] Wang X, Wei J, Schuurmans D, et al. Self-consistency improves chain of thought reasoning in language models. preprint [arXiv:220311171](https://arxiv.org/abs/2203.11171). 2022.
- [273] Kojima T, Gu SS, Reid M, et al. Large language models are zero-shot reasoners. preprint [arXiv:220511916](https://arxiv.org/abs/2205.11916). 2022.
- [274] Ahn M, Brohan A, Brown N, et al. Do as i can and not as i say: Grounding language in robotic affordances. preprint [arxiv:2204.01691](https://arxiv.org/abs/2204.01691). 2022.
- [275] Tanguchi T, Mochihashi D, Nagai T, et al. Survey on frontiers of language and robotics. *Adv Robot.* 2019;33(15-16):700–730. doi: [10.1080/01691864.2019.1632223](https://doi.org/10.1080/01691864.2019.1632223)
- [276] Steels L. The symbol grounding problem has been solved, so what's next? In: *Symbols and embodiment: debates on meaning and cognition*. Oxford: Oxford University Press; 2008. p. 223–244.
- [277] Taniguchi T, Ugur E, Hoffmann M, et al. Symbol emergence in cognitive developmental systems: a survey. *IEEE Trans Cognitive Develop Syst.* 2018;11(4):494–516. doi: [10.1109/TCDS](https://doi.org/10.1109/TCDS)
- [278] Taniguchi T, Nagai T, Nakamura T, et al. Symbol emergence in robotics: a survey. *Adv Robot.* 2016;30(11-12):706–728. doi: [10.1080/01691864.2016.1164622](https://doi.org/10.1080/01691864.2016.1164622)
- [279] Pezzulo G, Parr T, Cisek P et al. Generating Meaning: Active Inference and the Scope and Limits of Passive AI. *PsyArXiv*; 2023. Available from: psyarxiv.com/8xgzv.
- [280] Jung M, Matsumoto T, Tani J. Goal-directed behavior under variational predictive coding: dynamic organization of visual attention and working memory. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Macau: IEEE; 2019. p. 1040–1047.
- [281] Ito H, Yamamoto K, Mori H, et al. Efficient multitask learning with an embodied predictive model for door opening and entry with whole-body control. *Sci Rob.* 2022;7(65):eaax8177. doi: [10.1126/scirobotics.aax8177](https://doi.org/10.1126/scirobotics.aax8177)
- [282] Hoffmann M, Wang S, Outrata V, et al. Robot in the mirror: toward an embodied computational model of mirror self-recognition. *KI-Künstliche Intelligenz.* 2021;35(1):37–51. doi: [10.1007/s13218-020-00701-7](https://doi.org/10.1007/s13218-020-00701-7)
- [283] Juliani A, Arulkumaran K, Sasai S, et al. On the link between conscious function and general intelligence in humans and machines. preprint [arXiv:220405133](https://arxiv.org/abs/220405133). 2022.
- [284] Seth AK. Interoceptive inference, emotion, and the embodied self. *Trends Cogn Sci (Regul. Ed.)*. 2013;17(11):565–573. doi: [10.1016/j.tics.2013.09.007](https://doi.org/10.1016/j.tics.2013.09.007)
- [285] Bauermeister J, Lanillos P. The role of valence and meta-awareness in mirror self-recognition using hierarchical active inference. preprint [arXiv:220813213](https://arxiv.org/abs/220813213). 2022.
- [286] Hieida C, Nagai T. Survey and perspective on social emotions in robotics. *Adv Robot.* 2022;36(1-2):17–32. doi: [10.1080/01691864.2021.2012512](https://doi.org/10.1080/01691864.2021.2012512)
- [287] El Hafi L, Zheng Y, Shirouzu H, et al. Serket-SDE: a containerized software development environment for the symbol emergence in robotics toolkit. In: *IEEE/SICE International Symposium on System Integration (SII)*. Atlanta, GA: IEEE; 2023. p. 1–6.
- [288] Heins C, Millidge B, Demekas D, et al. pymdp: a python library for active inference in discrete state spaces. preprint [arXiv:220103904](https://arxiv.org/abs/220103904). 2022.
- [289] Reed S, Zolna K, Parisotto E, et al. A generalist agent. preprint [arXiv:220506175](https://arxiv.org/abs/220506175). 2022.
- [290] McGeer T. Passive dynamic walking. *Int J Robotics Res.* 1990;9(2):62–82. doi: [10.1177/027836499000900206](https://doi.org/10.1177/027836499000900206)
- [291] Brooks RA. Elephants don't play chess. *Rob Auton Syst.* 1990;6(1-2):3–15. doi: [10.1016/S0921-8890\(05\)80025-9](https://doi.org/10.1016/S0921-8890(05)80025-9)
- [292] Pfeifer R, Lungarella M, Iida F. Self-organization, embodiment, and biologically inspired robotics. *Science.* 2007;318(5853):1088–1093. doi: [10.1126/science.1145803](https://doi.org/10.1126/science.1145803)
- [293] Pfeifer R, Gómez G. Morphological computation—connecting brain, body, and environment. In: *Creating brain-like intelligence*. Berlin Heidelberg: Springer; 2009. p. 66–83.