A Marr-Inspired Framework for Raising “Good” Robots.-

HAMBURG, Sarah, JIMENEZ RODRIGUEZ, Alejandro, HTET, Aung and DI NUOVO, Alessandro <http://orcid.org/0000-0003-2677-2650>

A Marr-Inspired Framework for Raising "Good" Robots

*Sarah Hamburg, *Alejandro Jimenez-Rodriguez Aung Htet, Alessandro Di Nuovo,
*joint first authors

***Abstract*—Our current computer and AI systems are built on Neuroscience principles from almost a century ago. Recent advances in our understanding of biological computation have not crossed into computer science to catalyse advancements. We outline a multidimensional blueprint for a form of bio-inspired agents leveraging modern Neuroscience principles (including the co-localisation of memory and compute, plasticity, embodiment, active inference, and neurodevelopmental principles). We discuss how combining these core features could theoretically lead to cognitive agents that are aaligned to our prosocial values, transparent, explainable, and energy efficient (i.e., "good" robots). In particular, we leverage Marr's tri-level framework and advocate for an "Implementation Level" consisting of embodied neuromorphic hardware, an "Algorithmic Level" consisting of Active Inference, and a "Computational Level" consisting of prosocial goals (supported by evidence of prosociality catalysing the development of our own complex cognitive abilities). A developmental process scaffolds different prosocial computations over time. Supporting our perspective, we include simulation data demonstrating the transfer of priors between two different prosocial behaviours (Computational Level) via Active Inference (Algorithmic Level), supported by an embodied process (Implementation Level). Agent behaviour is transparent and explainable throughout. We advocate for this blueprint as a guide in creating capable, ethical, and sustainable machine intelligence.**

***Index Terms*—Neuro-developmental Robotics; Embodiment; Bio-inspired agents; Neuromorphic Computing; Active Inference; Machine intelligence**

I. Introduction: Closing the gap between Neuroscience and AI

CURRENT computer and AI systems predominantly leverage principles from the early dawn of neuroscience and psychology, almost a century ago. This includes von Neumann computer architecture (see Neuron Analogy section in Neumann's EDVAC First draft) [1] and the McCulloch-Pitts artificial neuron model [2]. The Hebbian learning principle of "cells that fire together, wire together" (1949) became foundational for unsupervised learning algorithms; the work of Neuroscientists Hubel Wiesel (1950-60s) was later adapted to form the principles behind convolutional neural networks; theories about learning and decision making from behavioural psychology formed the foundations of Reinforcement Learning (RL). Significantly, our understanding of biological cognition has experienced exponential growth since these early Neuroscience and Psychology discoveries, while Computer Science is yet to fully capitalise on such new knowledge and paradigm shifts.

Authors with Sheffield Robotics and the Department of Computing, Sheffield Hallam University, Sheffield, United Kingdom (e-mail: a.dinuovo@shu.ac.uk)

This work was supported by the U.K. Engineering and Physical Sciences Research Council (EPSRC) with the Grant EP/X018733/1 Autonomous Learning and Development in Embodied Neuromorphic Systems (ALDENS), and from the European Commission with the Horizon Europe Grant 101120727 Performance in Robot Interaction via Mental Imagery (PRIMI).

Manuscript received XXX; revised XXX.

Although current AI can sometimes appear intelligent it is criticised for a lack of understanding and absence of skills required for many real-world applications [3]. Key limitations include:

- Requirements for extensive training, computation, memory and energy
- Difficulties coping with (and exploiting) noise, variability, and uncertainty
- Difficulties generalising across tasks and environments
- Poor performance on tasks requiring embodied intelligence, such as adaptation to external conditions
- Inadequate transparency, interpretability, and explainability [4]

All mainstream AI approaches struggle with some degree of limitation across these areas, though the severity and nature of the challenges vary. Many AI researchers doubt that merely scaling-up current approaches will overcome such limitations, and instead argue for new inspiration from naturally intelligent systems [5]. Notable areas where contemporary AI falls short, but humans excel, include – learning with minimal data, abstract thinking (understanding concepts that are not immediately present or tangible), creativity, problem solving, adaptability, reasoning, decision making and collaboration [6]–[9]. Human cognition is also highly energy efficient. The estimated power consumption of the adult brain is 20 Watts per day [10] – equivalent to one energy efficient lightbulb.

We define a "good" robot as one that excels in such domains. The most pertinent of which for humankind likely relate to energy requirements and skills which may contribute to AI safety and alignment – including social and moral reasoning and collaboration.

According to the human self-domestication (HSD) hypothesis, which integrates anthropological insights with neuroscience, human brains evolved by selecting for traits associated with reduced aggression and increased cooperation, which in turn led to the emergence of complex languages and sophisticated cultures, including teaching and tool use [11]. Consequently, such processes may also be important in the context of developing AI that is both intelligent and aligned. In this sense, a "good robot" is a domesticated robot.

Zador et al. state that our challenge is to determine how best to exploit the synergies and overlaps in neuroscience and computational science to advance AI [5]. We argue that achieving this mission requires a multidimensional approach. Considering multiple dimensions is crucial for understanding and engineering complex systems – cognition that is both embodied and situated within a social and environmental context is one such system. A multidimensional approach enables interconnectedness and emergent behaviour to be explored and harnessed, and creates opportunities to optimise for ethical and social aspects alongside scientific and technical.

We aim to demonstrate the value of combining a minimal set of the following dimensions: bio-inspired "hardware" (specifically, embodied neuromorphic systems), "software" (the active inference framework), "training" (neurodevelopmental approach) and motivations (prosocial goals). This multidimensional approach can be framed within Marr's Tri-Level Hypothesis for information processing systems [12] – prosocial goals align with the computational level, active inference with the algorithmic level, and embodied neurodevelopmental systems with the implementation level. The neurodevelopmental approach is considered as a temporal dimension across all three levels (see Figure 1).

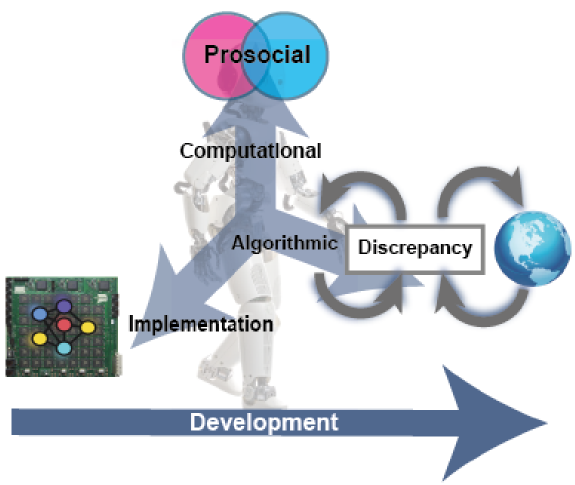

Fig. 1. A diagram illustrating the proposed multidimensional framework situated within Marr's Tri level hypothesis [12], with an additional temporal dimension as shown by the Development arrow (Neurodevelopmental approach to training occurring over time). To illustrate the implementation level, shown is an iCub humanoid robot and neuromorphic hardware (SpiNNaker). To illustrate the algorithmic level, shown is a diagram of the Active Inference framework [13]. Each dimension is discussedin detail in its respective section below. Our simulation study serves as a proof of concept to illustrate implementation of the framework

We postulate that this multidimensional framework, used as a design principle, may lead to the emergence of a novel form of cognitively advanced agents that are aligned to our prosocial values, explainable, and energy efficient. In line with a neurodevelopmental approach, these agents will be "raised" by humans. The following sections detail our blueprint for raising "good" robots (in both a capability and prosocial sense), for a collaborative future alongside machines. Although framed within the lens of Marr's Tri-Level Hypothesis, all dimensions are proposed as synergistically interacting (see section Inter-Dimensional Relationships).

II. The Implementation Level: Embodied Neurodevelopmental Systems

Some of the limitations of current AI are due to mainstream computing architectures. This includes an unnatural discretisation of time imposed by mainstream processing and sensing architectures [14]. Furthermore, the Von Neumann architecture separates memory from instructions, creating a bottleneck that is not present in biological brains [15]. Indeed, while modern computers have extensive storage well beyond human capacities, they are inefficient as they usually necessitate specific commands to retrieve particular data stored in their memory banks. In contrast, human memory functions via associative networks, enabling spontaneous recall triggered by contextual cues or patterns, as opposed to relying on precise queries.

Marr's implementation level relates to the physical realisation of computation [12]. Applied to robotics, this would encompass how cognitive processes are dynamically instantiated in a body, circuits, and sensors, and in turn how this instantiation shapes cognitive processes. As mentioned above, dimensions are synergistically interacting, therefore implementation level constraints and signals will impact algorithmic level priors for example, as discussed in the Inter-Dimensional Relationships section.

A. Neuromorphic Systems

Neuromorphic computers are inspired by the brain's structure and function. They aim to be more powerful and energy efficient than traditional computers. Key features include real-time parallel processing, event-driven local computation, plasticity, and low power consumption. The resulting promise is for highly efficient, adaptable systems, capable of real-time multimodal integration. For example, a robotics study by Tang et al. demonstrated 100 times less energy consumption and comparable accuracy to traditional methods using a neuromorphic algorithm (spiking neural network; SNN) for simultaneous localisation and mapping (SLAM) [16].

Examples of currently available neuromorphic computers include the Human Brain Project's SpiNNaker 2 [17] and BrainScales 2 [18]. The world's first supercomputer capable of simulating networks at the scale of the human brain is neuromorphic and due to become operational in 2024 (DeepSouth). Importantly, neuromorphic computers can be combined with neuromorphic sensors (such as event cameras) and embedded within a robot to form embodied neuromorphic systems. Notably, neuromorphic hardware has been described as a "key enabling technology for the development of a unique generation of autonomous agents endowed with embodied neuromorphic intelligence" [19].

An alternative form of bio-inspired hardware is biomolecular computing, which exploits the properties of organic materials and molecules for computation. Such hardware could offer similar advantages to neuromorphic systems in terms of parallelism and low energy-consumption [20], in addition to the possibilities of self-assembly and repair [21]. However, scaling such systems remains a challenge. It seems likely that in the future a biohybrid-neuromorphic approach, such as the use of cerebral organoids [22], may prove a useful implementation level substrate.

B. Embodiment

It has been argued that one major limitation of mainstream machine learning is the absence of a body to support self-determined learning via autonomous interaction with the environment [23].

Neuropsychology has extensively evidenced the embodied and situated nature of human cognitive abilities, which are formed not only within the brain but are shaped by the body and the experiences acquired through it during development (i.e. motor movements and interactions with objects and people). There is particularly strong evidence for the role of embodiment in the acquisition of perceptual, language, social and numerical skills [24]–[27]. Emulating embodied learning mechanisms in artificial agents may therefore promote the acquisition of such advanced skills.

This Neurodevelopmental approach aligns with the 1950 hypothesis by Alan Turing, "Instead of trying to produce a program to simulate the adult mind, why not rather try to produce one which simulates the child's brain? If this were then subjected to an appropriate course of education one would obtain the adult brain" [28].

C. Neurodevelopmental Approach

The Neurodevelopmental approach to robotics (or "Developmental Neurorobotics") is an interdisciplinary research paradigm combining computational modelling, developmental psychology and robotics to realise an embodied artificial intelligence [29].

The approach lends heavily from Piaget's theories of cognitive development, including an emphasis on self-determined learning via interaction with the environment [30]. Embodied agents build models (i.e, learn) based on their own interactions with the world, rather than rely on pre-trained models.

According to Piaget [30], key features of learning and development include:

- learning is cumulative and progresses in complexity
- concrete and abstract concepts are a continuum; both are learned by linking concepts to embodied perceptions [31]
- learning results from self-exploration with the world; often in combination with social interaction
- the importance of sensorimotor skills (including the discovery of one's own body), linguistic skills, and social skills

These themes are important in Developmental Neurorobotics for realising embodied artificial agents. Implementing this learning (or "training") framework in neuromorphic systems is enabled though the above outlined features of neuromorphic computers – including real-time parallel processing, event-driven local computation, multimodal integration, and plasticity. For example, Spike Timing Dependent Plasticity (STDP) provides a biomimetic method for implementing learning over time – starting with simple synaptic modifications in response to temporal relationships between events, which are in turn refined over longer periods, producing changes in network architectures (mimicking neurodevelopmental processes, such as synaptic pruning). Accordingly, learning occurs dynamically over the lifetime of the agent in response to embodied interactions within its specific environment and "curriculum" experienced, rather than via pre-training during development.

Vygotsky's Sociocultural Theory of Cognitive Development [32] is also important in informing the Neurodevelopmental approach. This theory emphasises the role of social interaction and the cultural environment in mental abilities. Vygotsky argued that higher mental functions could only develop through interaction with more advanced peers. He also postulated the existence of the Zone of Proximal Development (the difference between what a learner can do without help and what they can do with help), which forms the theoretical underpinnings of "scaffolding" in learning.

This Neurodevelopmental approach can be paralleled with Brooks "Behaviour-based robotics" framework [33], which emphasises the creation of autonomous robots with behaviours that are added incrementally ("incremental complexity"). Within this behaviour-based framework, foundational behaviours scaffold more sophisticated behaviours – similar to developmental stages. Practically, these theories can be leveraged in the design of effective educational curricula for agents, in order to train agents in accordance with the neurodevelopmental approach (i.e. "raise" agents with required qualities).

Marr's framework has been criticised for not sufficiently addressing temporal aspects of cognitive functions. The Neurodevelopmental approach overcomes this limitation.

III. The Algorithmic Level: Active Inference

Marr's algorithmic level [12] relates to processes employed in transforming inputs into outputs. Applied to robotics, this would encompass the models used to process information and perform actions.

A. Overview of the Active Inference Framework

Originating from Neuroscience, the Active Inference Framework (AIF) offers a biologically plausible and unified explanation for how the brain processes information, learns, and generates behaviour [34]. This includes solving "hard exploration problems" and accounting for mechanisms of natural agency and behaviour [35]. We aim here to provide a high-level overview of AIF relevant to Embodied Neurodevelopmental Systems. For an in-depth explanation and discussion of AIF see Parr et al. [13]. For an in-depth discussion of AIF for learning and development in embodied neuromorphic agents see Hamburg et al., 2024.

Within AIF, the brain models the world as a set of probabilities which it uses to make inferences and predictions about the world. The brain actively works to minimise "surprise" by aligning predictions and observations as closely as possible ("surprise" is a measure of uncertainty about the world, considering the quality of data. It is also often referred to as "prediction error" or "free energy"; the later provides an upper bound to surprise.

Free energy (and hence surprise) minimisation is achieved by 1) adjusting models (i.e. altering perception), and/or 2) selecting actions that maximise information gain and minimise prediction errors (e.g. turning your head towards an unknown noise). Consequently, AIF is said to provide a dual account of both perception and action.

AIF involves a continual loop of prediction, perception, and action. Over short-timescales, perception optimises beliefs about the world. Over long-timescales, learning optimises beliefs about the relationships between the variables that constitute the world [36]. Those processes occur through the minimisation of variational and expected "free energy" respectively. The ability to account for both short- and long-term learning is particularly advantageous for neurodevelopmental frameworks, which operate over and integrate these different timescales in learning.

Compared to RL, AIF offers a more integrated view of perception and action, along with more flexible goal-setting based on prior preferences. In RL, the reward function defines an agent's goal and allows it to learn how to best act within the environment to maximize expected reward. In AIF, any type of outcome may be more or less preferred – the implicit reward is a feature of the agent, not the environment it inhabits [37].

AIF bypasses problems associated with defining reward functions (which can be difficult, particularly for real-world tasks [35]) and instead replaces these with prior beliefs about preferred outcomes. Agents learn their own prior preferences and goals are flexible. Consequently, AIF extends RL, encourages exploration and information seeking, and equips agents with intrinsic curiosity [38], [39].

While AIF is rooted in Bayesian inference, and the two approaches share a probabilistic modelling approach, Bayesian models typically treat action as a separate stage of processing (decisions are made based on inferred probabilities), while in AIF action is treated as an integral part of perception (the brain actively seeks out sensory information that reduces uncertainty and prediction error).

AIF shares some similarities with perceptual control theory (PCT) [40], however in AIF action control has anticipatory/feedforward aspects (based on generative models), while in PCT it is assumed that feedback mechanisms are sufficient for action control. Furthermore, in AIF, motivational processes can modulate the contribution (i.e. weighting) of different goals in action control (for further discussion see Parr et al) [13]. Together this suggests AIF if capable of entailing a richer hierarchical architecture.

B. Theoretical and Empirical Support

Support for the value of AIF at the "algorithmic level" of Embodied Neurodevelopmental Systems is both theoretical and empirical. Theoretically, embodiment is a key feature of AIF – perception and cognition are deeply situated and intertwined in the embedded context of the agent and its environment [41]. In AIF, the brain has even been described as "taking a back seat to the body" [41]. There is also an emphasis on brain-body-environment interactions – in AIF, there is no distinction between "agent" and "environment". Instead, Markov blankets (conceptual boundaries that isolate sets of variables) are employed to represent boundaries between systems with an external and internal state.

The drive to reduce uncertainty underpinning AIF has been described as comparable to curiosity [36]. The emphasis on curiosity-driven embodied behaviour in AIF suggests it may offer a useful framework for self-supervised learning, as in neurodevelopmental frameworks.

Further benefits may include the integration of multiple sensory streams, learning from sparse and noisy observations, transparency and explainability [42]. Regarding explainability, AIF offers a set of prior beliefs about decisions that represent explanations for behaviour, and AIF systems that appear to understand their actions have been demonstrated [43]. This contrasts with current "black box" machine learning methods. AIF could potentially generate high-order cognitive and metacognitive capabilities, such as monitoring, self-explainability and in some degree "awareness" [38]. Consequently, it has been suggested that AIF may lead to embodied agents that are context adaptive, safe, social, and collaborative [42].

Empirically, AIF has been shown to perform as well as traditional ML methods in simple environments, and better in environments featuring volatility, ambiguity and context sensitivity [42]. AIF appears particularly useful for applications where the dynamics of the robot and/or the task are uncertain [38]. Implementations of AIF in embodied systems have included simulated robot arms for searching, reaching and manipulating [44]–[46]; a model for estimation and control in a humanoid robot [38], [47]; and multisensory body perception and adaptive control (action) in a humanoid robot [48]. Further skills demonstrated include navigation [49], [50], fault-tolerant behaviour [51], and complex social cognition [52]. Learning has included the ability to generalise prior knowledge to new stimuli, resulting in a "one-shot learning" capacity qualitatively similar to that observed in humans [36].

In neuromorphic systems, AIF was recently shown to naturally yield Hebbian plasticity [53]. The authors suggest this approach may dramatically reduce the complexity of designing self-learning systems. Gandolfi et al. [54] also recently demonstrated plasticity and rapid unsupervised learning in a neuromorphic system using AIF principles. The authors suggest their experiments could be adopted to implement brain-like predictive capabilities in neuromorphic robotic systems. Furthermore, an "embodied" neuromorphic AIF system ("DishBrain") recently demonstrated rapid apparent learning of the computer game Pong [55]. The authors claimed the system exhibited "synthetic biological intelligence". These examples suggest there may be important opportunities for interactions between the "algorithmic level" (active inference) and "implementation level" (neuromorphic hardware) (see section Inter-Dimensional Relationships).

In the field of neurorobotics it has been said, "A real breakthrough in the field will happen if the whole system design is based on biological computational principles, with a tight interplay between the estimation of the surroundings and the robot's own state, and decision making, planning and action" [19]. We suggest that AIF is well placed to meet these requirements through the "algorithmic level" [12] of Embodied Neurodevelopmental Systems.

IV. The Computational Level: Prosocial Goals

Marr's computational level (1982) [12] relates to the tasks or goals a system is trying to achieve. The term "computational level" here refers to the highest level of abstraction of information processing systems and is not to be confused with algorithmic computations (covered within the algorithmic level, detailed above).

The human self-domestication hypothesis (HSD) posits that our unique set of human traits (including complex languages, extensive tool use, and sophisticated cultures) has emerged from an evolutionary process of self-induced domestication, through which humans evolved to be less aggressive and more cooperative [56].

According to HSD, evolution in the middle and late Paleolithic was characterised by selective pressures for less aggressive partners, resulting in more prosocial individuals. Prosocial behaviour is indented to benefit others, or at least promote harmonious relations [57]. Examples include donating, sharing, helping and cooperating. Prosocial behaviour and intelligence are both genetically influenced, and extensive research demonstrates a positive association between these two traits [58], [59].

The HSD posits that more prosocial individuals led to an increase in social contacts and complex community structures – in turn leading to more sophisticated teaching, learning, and linguistic abilities [11]. Indeed, the process of domestication has been directly linked to vocal learning, with domesticated animals typically display more complex vocalisations compared to wild relatives [11]. A key outcome of HSD is a prolonged developmental window and parental care – enhancing learning opportunities via exposure, imitation and culture (vs. innate knowledge), in turn enhancing the acquisition of complex skills and behaviours [11].

Further, Barrett et al. [60] highlighted how social and technical skills interact in mutually reinforcing ways, and posit that human cognition is a cultural artefact: "In a reversal of the standard view, language and other complex cognitive skills do not form the underpinnings of our sophisticated material cultures but are, instead, considered to be their manufactured products".

In accordance with evidence that prosocial cognition and behaviour catalysed the evolution of our own complex cognitive abilities, we advocate for prosocial goals forming the "computational level" of Embodied Neurodevelopmental Systems. Practically, prosocial goals might entail a suite of available actions (or behavioural constraints) which benefit other agents or organisms. For example, in our simulation of rat behaviour (below), rats are able to choose whether or not they provide comfort to another rat experiencing discomfort.

In current AI systems, objective functions often optimise for individual agent utility. This may promote behaviour that is self-serving rather than prosocial, and indeed may diverge from prosocial goals in an effort to maximise rewards. In current approaches, alignment behaviours must also be explicitly programmed and do not necessarily adapt and evolve. Forming a computational level of prosocial goals may lead to agents prioritising cooperation with humans over competition (and exploitation). Such prosocial behaviours may adapt and evolve over time.

V. Inter-Dimensional Relationships

Although we have considered each dimension in turn, these are not isolated and instead synergistically interact (see Figure 2).

For example, AIF processes are dependent upon preferred states. Prosocial goals at the Computational Level could entail a mechanism through which such Algorithmic Level priors are furnished. In turn, AIF processes are fundamental for updating such higher-level goals and choosing aligned actions. A key challenge for AIF remains the design of meaningful prior beliefs (i.e. preferences) – our architectural addition of a Computational Level consisting of prosocial goals could help overcome this.

Furthermore, priors and available actions are dependent upon Implementation Level constraints (i.e. embodied neurodevelopmental systems). Such constrains are integral to optimising AIF. In turn, AIF could modify and optimise Implementation Level plastic neuromorphic systems (e.g. at the level of the synapse [61]).

The potential value in combining AIF and neuromorphic circuits should not be underestimated. Indeed, AIF modelling processes reproduce a range of neural phenomena (e.g. theta-gamma coupling, place-cell activity) and related agent behaviours (e.g. reward seeking, context learning) in traditional computing systems [62]. As neuromorphic architectures are the natural substrate for AIF processes, implementing AIF within such systems may generate novel computing tools.

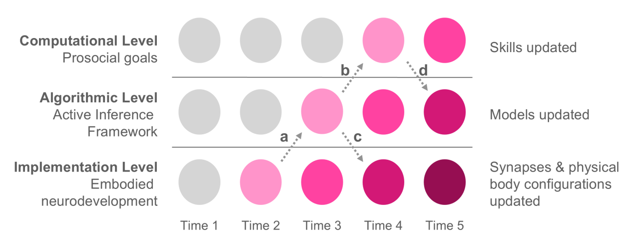

Fig. 2. A diagram illustrating example interactions among the multiple dimensions of an agent over time. The agent is illustrated by the grey and pink circles; each row represents a level of the framework; each column represents the agent at a particular time point; for example, the agent at time 1 is illustrated by the three circles in this column, wherein the top circle represents the agents computational level at time 1, the middle circle represents the agent's algorithmic level at time 1, and the bottom circle represents the agent's implementation level at time 1. Labelled arrows provide examples – a) represents available actions (e.g. STDP and actuator rules) and constraints for AIF; b) represents AIF scaffolding the learning of prosocial skills; c) represents AIF performing actions on the agent's physical implementation (e.g. synaptic connections, actuator positions); d) represents prosocial goals providing information about preferred states for AIF. Colour gradient change represents increasing complexity over time.

Further, computational intractability is a key criticism of Bayesian predictive processing approaches, including AIF.

Kwisthout & van Rooij [63] recently leveraged computational complexity theory to investigate this issue and conceptualised subcomputations within Marr's Computational Level (independent of Algorithmic or Implementation Level processes). Importantly, the authors demonstrated the necessity of constraints on subcomputations in circumventing computational intractability. Specifically, computations can be performed tractably when the topological structure of the Bayesian network is constrained, when each variable can take a small number of distinct values, and when the search space of possible predictions and hypotheses is small. We suggest that situating AIF at the Algorithmic Level, as opposed to the Computational Level, will provide both bottom-up and top-down constraints to enhance tractability in line with these findings.

VI. Raising Good Robots: Simulation Study

We have simulated the behaviour of rats as an example of the potential behaviour of robots. In various animal species, including rats, juvenile social thermoregulation via huddling is thought to confer later social and altruistic behaviours in adulthood (e.g. social ("filial") huddling and contact-comfort) [64].

Our model is composed of two prosocial goals (filial huddling and contact-comfort), each structured as a partially observable Markov decision process (POMDP). An underlying developmental process (also structured as a POMDP) selects the available actions or policies depending on internal cues from the agent (see Figure 3). Simulations were performed using python3 and PYPMDP package [65]. Full models are provided in the appendix and code is available on Github.

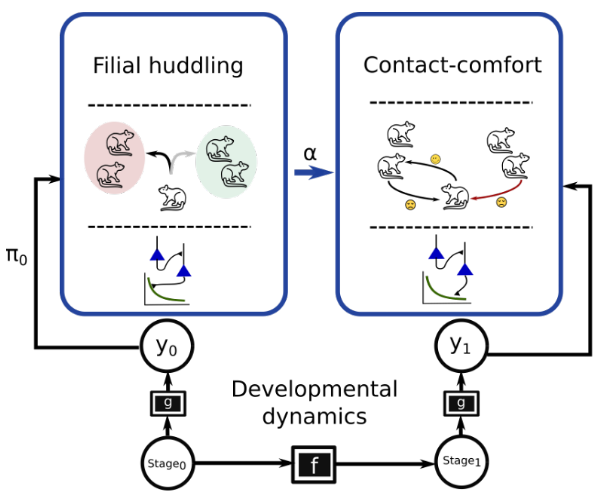

Fig. 3. Blue boxes illustrate the different dimensions of the simulation – Implementation Level (bottom: embodied developmental; change in BAT), Algorithmic Level (middle: POMDPs), and Computational Level (top: huddling in pups/comforting in adults); f is the developmental transition function; g is the function providing the likelihood of observation (y); π is the action (huddle/comfort); α is the social prior that is transferred across developmental timepoints labelled as stages 0-1.

A. Developmental POMDP

We used a POMDP to model a developmental process as active inference – where the thermogenesis in brown adipose tissue (BAT) is used as a proxy for two different developmental stages which we assume as discrete for simplicity [64]. The BAT thermogenesis signal is therefore the environment of the Developmental POMDP, modelled as:

$$bat(t) = e^{-\beta t} + \mathcal{N}(0, \sigma^2) \tag{1}$$

β defines the time scale of development and, effectively, the duration of each of the developmental stages, σ is the variability of the physiological process. This developmental signal is observed by the developmental process:

$$o(t) = \Theta(bat(t) - \pi) \tag{2}$$

Where $\Theta(\cdot)$ is the Heaviside function, and π is an arbitrary threshold for each developmental stage. Developmental and huddling/comfort timescales were kept equal for simplicity. At each step, the agent computes an approximate (updated) belief about the current developmental stage by minimizing the variational free energy (VFE) and approximating the posterior distribution:

$$q(s) \approx p(s|o) \tag{3}$$

The developmental process is thus purely inferential as actions performed through downstream policies do not affect the agent's internal model about its developmental stage. However, more elaborate developmental processes are possible.

B. Filial huddling POMDP

We model the action selection of filial huddling as minimising expected free energy (EFE). Following Wilson [64], we define an abstract huddling process as follows: We start with a set of rat pups which are assigned to either of two huddles $pup_i \in \{1, 2\}$. At each step of the simulation, a random pup is selected and the agent is either assigned to that pup's huddle with probability p or stays in its own huddle with probability $1 - p$, where:

$$p = (1 - e^{-T_\alpha})^{-1} \tag{4}$$

Here, $T_a = 2\alpha T$ is a temperature parameter provided by the developmental process (see above) that increases the probability of creating huddles. The agent stores a belief about which huddle it belongs to $q(s) \approx p(s|o)$ given the observation of what group it has been assigned. As a result of the EFE, the agent decides to either have contact or not with its new litter mates. From this action, the model collects observations about the associations of the agent with either huddle (α is the association strength with the current huddle). The strength of association is increased by:

$$\Delta\alpha_i = \gamma(\pi_0 - \sum_{i \neq k} \alpha_k) \tag{5}$$

Notice that this implies that the strength will decrease with no contact [64].

C. Altruistic behaviour POMDP

Altruistic behaviour is modelled through a contact-comfort task [66]. The same pups of the huddling simulation can now be in either of two states as adults: discomfort (0) and comfort (1). These states generate observations $o(t) \in \{cry, squeak\}$.

At each step of the simulation, a random adult is chosen to transition to discomfort and a random adult is chosen to be observed. The adult agent stores beliefs $q(s) \approx p(s|o)$ about whether each of the groups (huddles; H) is happy or sad $s_i \in \{H_1\ sad, H_1\ happy, H_2\ sad, H_2\ happy\}$ estimated by minimizing the VFE. The adult agent also possesses some priors over observations that are set by the developmental process and proportional to the associations generated during huddling:

$$P(H_i\ squeak) = \alpha_i \tag{6}$$

As a result of the minimization of the EFE, the adult agent decides whether to comfort or not the observed adult:

$$pup_i = \begin{cases} 1 & u^* = comfort \\ 0 & otherwise \end{cases} \tag{7}$$

where

$$u^* = \underset{u}{\operatorname{argmin}}\, EFE(o, u)$$

D. Implementation level details

As mentioned in section II, the implementation level is composed of embodiment, illustrated by the physiological signal simulated in this study, learning processes across development, illustrated by the learned affinities between agents, and a neuromorphic implementation that is possible thanks to recent process theories of AIF [62] and explicit free energy minimizing neural networks [53].

E. Simulation Results

In the early stages of development, the huddling policies are selected by the developmental process (Figure 4, left), which in turn generate filial associations with one of the huddles (Figure 4, right). In turn, those associations bias the consolation behaviours in later developmental stages (Figure). Notice that, in the absence of those priors, the agent does not discriminate between huddles with its consolation behaviour.

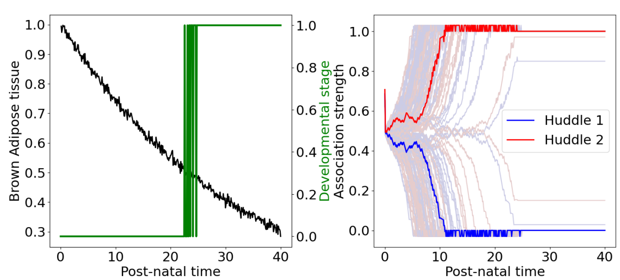

Fig. 4. Left: BAT signal and the corresponding beliefs about the developmental stage of the agent. Right: Filial huddling associations' evolution during the first stage of development. Notice sometimes the transition to a preferred group fails in the given timeframe – replicating natural variation in social skills. All simulations have 100 trials with 400 timesteps.

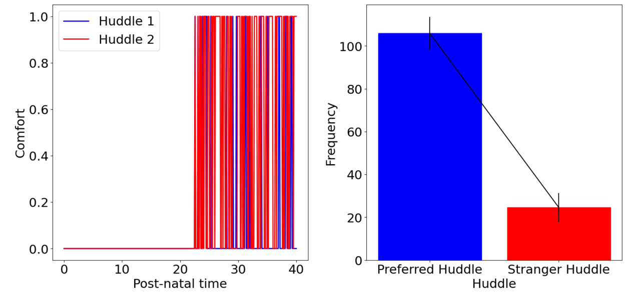

Fig. 5. Left: Comfort states in the two huddles are observed in the second stage of development (notice that policies and observations are not available during the first stage). Each line shows whether the given huddle is in discomfort or sad (0) or comfortable or happy (1). At each step, the transition to comfort is given by the actions of the agent. The transition to discomfort is the result of a random process. Right: Proportion of the total time each group spends in a comfort state, showing agents' clear preference for the prior associated huddle.

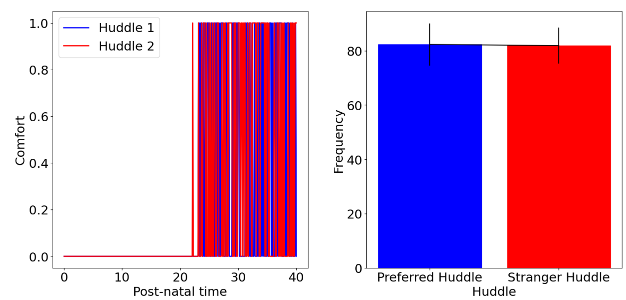

Fig. 6. As Figure 5 but without priors from huddling – notice that without these the agent spends equal time comforting both groups.

Our results demonstrate the transfer of priors from an embodied developmental processes (Implementation Level) to later prosocial behaviour (Computational Level) via AIF (Algorithmic Level). We additionally demonstrate the use of AIF to model developmental processes – including developmental stages, whereby new actions become available to an agent once an internal threshold is reached. Future work should explore potential cognitive advantages conferred by such social associations and behaviours (e.g. imitation learning), in addition to neuromorphic implementation through the use of Active inference's process theory.

Significantly, agent behaviour is transparent and explainable at every stage through biases specified by earlier developmental experiences. In this particular example, the parameter alpha can be seen as an explanation of the decision making process in the next developmental stage. This feature offers a significant advantage over current "black box" methods, wherein it is difficult to determine why an agent makes a particular decision. Consequently, the present approach could catalyse the use of AI in regulated and high-stakes domains, such as finance, healthcare and autonomous vehicles. Such accessible and interpretable parameters may also be exploited by the agent itself to confer metacognitive abilities (i.e. awareness and understanding of it's own thought processes), in addition to exploitation by other agents, and humans, to catalyse collective intelligence and shared decision making.

VII. Summary

Our novel bio-inspired architecture for embodied artificial agents draws upon interdisciplinary research to address longstanding limitations and demonstrates a pathway creating agents with advanced capabilities, specifically engineered to collaborate alongside us transparently and sustainably. We advocate for future research to further explore and validate this blueprint, particularly in larger scale implementations.

VIII. Appendix
POMDP formulation

We present a formulation of the different stages of development as well as the behaviours specific to each stage as a hierarchical Partially Observable Markov Decision Process (POMDPs) in the framework of discrete active inference [13] (figure 7).

We assume the agent goes through a sequence of developmental stages $\aleph_1, \aleph_2, \ldots, \aleph_n$. Each of the stages correspond to a particular POMDP, $\aleph_i = (S_i, O_i, X_i, P_i, \pi_i)$, where S_i is some finite space of hidden state spaces. O_i is a finite set of outcomes or observations. Note that there could be multiple factors S_i^m or sensory modalities O_i^k. X_i is finite set of control states or actions for that stage.

For each stage, the agent has a generative model P_i and a policy (or set of policies) that we consider fixed for this paper. The generative model is composed of a likelihood density $p(o_i^\tau|s_i^t) = \text{Cat}(\mathbf{A}_i)$, a transition probability $p(s_i^{\tau+1}|s_i^\tau, \pi_i) = \text{Cat}(\mathbf{B}_i)$, and priors over states $p(s_i^1) = \text{Cat}(\mathbf{D}_i)$ and a prior or preference over outcomes $p(o_i|C) = \text{Cat}(\mathbf{C}_i)$.

The transition between stages is driven by a developmental POMDP that uses the internal physiological state, and other clues from the environment to infer the developmental stage the agent is in. The space of outcomes is then $\mathcal{O} = O_1 \times O_2 \times \ldots O_n \times I$, where I is the space of internal outcomes observed from the body. We give a concrete specification in the following section. Finally, the states of the developmental process correspond to the different developmental stages.

Communication between stages happens by means of parameters α_i and β_i, which are learnable through the behaviour at each stage, such that, $\mathbf{C}_i = \alpha_{i-1}$, defining the preferences for the next stage, and $\mathbf{D}_i = \beta_{i-1}$, defining the beliefs or priors over hidden states.

A. Developmental Model

As mentioned in the main test, the generative process associated with development is a physiological signal that has two main functions. First, to set the time-scale of the transitions between the different stages, and second, to drive the transition between them activating the different policies and outcome modalities available at each stage. Note that in more complex model, there could be multiple signal and external cues that could influence this process. In this particular model, we have chosen a proxy for the Brown Adipose Tissue thermogenesis [64]. This signal is given by:

$$bat(\tau) = e^{-\beta\tau} + \mathcal{N}(0, \sigma^2) \tag{8}$$

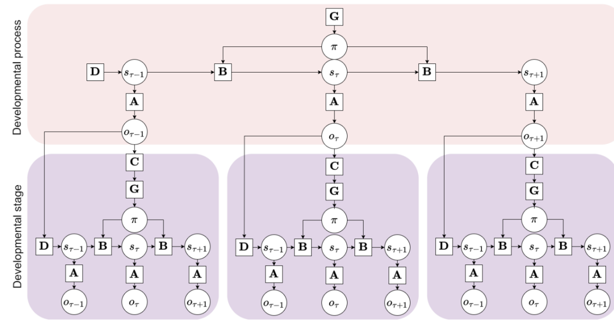

Fig. 7. Hierarchical POMDP (factor graph) for the multidimensional Marr-inspired framework implementation presented in this paper. At level one, the specific behaviours of the different developmental stages are unfolded by minimizing the expected free energy (Details in the text). At level 2, the developmental process modulates the policies, states and observations from each developmental stage by marginalizing the different categorical distributions involved.

With $\sigma = 0.1$. From this equation, the agent derives an observation $o_\tau = \Theta(bat(\tau) - 0.5)$. Note that all the model will have a unique time because the difference in timescales is given by the parameter β. In order to define the generative model [13], [67] we define the following categorical distributions:

$P(o_\tau|s_\tau) = Cat(A)$, the likelihood of observed signals given a particular developmental stage, $P(s_{\tau+1}|s_\tau, \pi) = Cat(B)$, the transition function between developmental stages, $P(o_\tau|C) = Cat(C)$, the preference over outcomes and the preference over stages, $P(s_\tau) = Cat(D)$; for states $s_\tau \in \{stage_1, stage_2\}$ and observations $o_\tau \in \{0, 1\}$.

The dynamics in this case is trivial, as we do not have additional sources of uncertainty apart from the noisy BAT signal. The previous distributions are parametrized by the following matrices in our model:

$$A = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \tag{9}$$

Where we have given some uncertainty to the current observation given a particular stage (BAT thermogenesis continues despite being replace by other processes). The transition tensor is parametrized by the different behaviours expressed by the level below ($behavior_i$ is the behaviour associated to the i^th developmental stage). Note that these behaviours could possibly affect the developmental dynamics by affecting the underlying physiological process, however, they do not in our current model.

$$B(:,:,behavior_0) = \begin{pmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{pmatrix} \tag{10}$$

$$B(:,:,behavior_1) = \begin{pmatrix} 0.5 & 0.1 \\ 0.5 & 0.9 \end{pmatrix} \tag{11}$$

The transitions in the second column for the first matrix are not defined (i.e. the agent being in the second developmental stage and expressing behaviours from the first one). The matrices C and D are uniform.

The minimization of the free energy [67] gives the current developmental stage s_τ and the corresponding action $u \in$

$\{\textit{Filial huddling}, \textit{Altruistic behavior}\}$, which is the behaviour to be expressed in the level below.

The behaviours expressed at the different developmental stages are also modelled as a POMDP. They can be seen as a single model with disjoint sensory modalities, states and policies. In their more general form, the categorical distributions can be defined as:

$A_i \in R^{m\times n\times k}$: A set of matrices defined for each income modality i of which, some of them belong to the first development stage, some to the second and so on (i.e. $i \in \{\{1\},\{2,3\}\}$). Each of the matrices have m possible outcomes, n states and k state factors. The state factors are also partitioned for the different stages.

Similarly, $B_i \in R^{m\times m\times k}$ (we reuse the indexes for simplicity) defines a set of transition matrices, one for each state factor with different state factors associated to different stages (in the first stage, the states represent the huddle the agent belongs to, in the second, the beliefs about the comfort/discomfort of the members of each group). Each matrix has, m possible states and k actions that are also partitioned for the different stages.

The developmental modulation of the different stages by the developmental process, can be done by selecting the appropriate matrices for the selection of actions and the update of beliefs, by means of providing appropriate priors, C over modalities, D over states and E over policies.

Even though we think it is useful to see the behaviour of the agent in this unified way, the different generative models for the two stages are effectively disjoint and we will present them now as two separate POMDPs to avoid excessive indexes.

B. Filial Huddling

As described in the main text, we start with a set of pups which are assigned to either of two huddles $pup_i \in \{1,2\}$. At each step of the simulation, a random pup_i is selected and the agent is either assigned to huddle pup_i with probability p or the other huddle with probability $1-p$, where:.

$$p = (1-e^{-T_a})^{-1} \tag{12}$$

Here, $T_a = 2\alpha T$ is a temperature parameter provided by the developmental process that increases the probability of creating huddles, and α is the association strength with the current huddle. The huddle assignment becomes the observation $o_\tau \in \{\textit{huddle1}, \textit{huddle2}\}$. The agent can be in one of two states that represent the huddle it belongs to $s_\tau \in \{\textit{huddle1}, \textit{huddle2}\}$, and the actions available are $a_\tau \in \{\textit{Contact}, \textit{No contact}\}$. We assume that making contact with the other pups in the assigned huddle increases the association with that huddle and that the decision of making contact depends on the belief of the agent about which huddle it belongs to.

We use the VFE to approximate belief about huddle belonging $q(s) \approx$ given the observation of what group it has been assigned to. Similar to the developmental model, we parametrize the likelihood by the matrix:

$$A = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \tag{13}$$

capturing the expectation that the assignment matches with the huddle the agent belongs to. The transition probabilities are given by:

$$B(:,:,\textit{Contact}) = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \tag{14}$$

$$B(:,:,\textit{No Contact}) = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \tag{15}$$

In this case, the decision of No Contact, could generate a transition in the internal state of the agent. The strength of association is increased by:

$$\Delta\alpha_i = \gamma(\pi_0 - \sum_{i\neq k}\alpha_k) \tag{16}$$

Where $\pi_0 = 1$ when contact is made and $\pi_0 = 0$ otherwise. Notice that this implies that the strength will decrease when no contact is made [64].

C. Altruistic behaviour POMDP

Altruistic behaviour is modelled through a comfort-discomfort task. The same pups of the huddling simulation now can be in either of two states: comfortable (1) or in discomfort (0). We define two outcome modalities: $o\{1\}_\tau \in \{\textit{cry}, \textit{squeak}\}$ and $o\{2\}_\tau \in \{\textit{Huddle}_1, \textit{Huddle}_2\}$; and two state factors $s\{1\}_\tau \in \{\textit{sad}, \textit{happy}\}$ and $s\{2\}_\tau \in \{\textit{Huddle}_1, \textit{Huddle}_2\}$. At each step of the simulation, a random pup is chosen to transition to discomfort and a random pup is chosen to be observed. The agent computes beliefs $q(s) \approx p(s|o)$ about whether each of the groups is happy or sad, estimated by minimizing the VFE. The likelihood functions are encoded in the block matrix:

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \tag{17}$$

with

$$A_1 = A_2 = \begin{pmatrix} 0.3 & 0.7 \\ 0.7 & 0.3 \end{pmatrix} \tag{18}$$

Which encodes the fact that, by default, a cry or a squeak are evidence of either of the groups being sad or happy. The transition matrices between states, are also encoded in block matrices that encode the difference state factors:

$$B(:,:,\textit{Help 1}) = \begin{pmatrix} B_1 & 0 \\ 0 & I \end{pmatrix} \tag{19}$$

$$B(:,:,\textit{Help 2}) = \begin{pmatrix} I & 0 \\ 0 & B_1 \end{pmatrix} \tag{20}$$

$$B(:,:,\textit{No Help}) = \begin{pmatrix} B_2 & 0 \\ 0 & B_2 \end{pmatrix} \tag{21}$$

with

$$B_1 = \begin{pmatrix} 0.9 & 0.9 \\ 0.1 & 0.1 \end{pmatrix} \tag{22}$$

$$I_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{23}$$

and

$$B_2 = \begin{pmatrix} 0.1 & 0.1 \\ 0.9 & 0.9 \end{pmatrix} \tag{24}$$

The transitions are defined for each of the available actions $a_\tau \in \{\textit{Help 1}, \textit{Help 2}, \textit{No Help}\}$. Notice that that, given help, their beliefs about the happiness state factor tends to transition to happiness irrespective of the group the agent belongs to; the agent is altruistic by default. However, when there is no help, the belief about both groups tends to change about "sadness". As an influence from the previous developmental stage, we define a prior over observations given by the associations generated during huddling:

$$P(o_\tau|C) = Cat(C) \tag{25}$$

where the matrix C is selected such that $P(H_i\ \textit{squeak}) = \alpha_i$, where α_i is the association strength created during the previous stage. Finally, as a result of the minimization of the EFE, the agent decides whether to comfort or not the observed pup:

$$pup_i = \begin{cases} 1 & u^* = \textit{comfort} \\ pup_i & \textit{otherwise} \end{cases} \tag{26}$$

where

$$u^* = \operatorname*{argmin}_u \textit{EFE}(o, u)$$

. That is, the result of an action is a transition of the state of the given huddle to comfort whenever he action is to help.

REFERENCES

[1] M. D. Godfrey and D. F. Hendry, "The computer as von neumann planned it," *IEEE Annals of the History of Computing*, vol. 15, no. 1, pp. 11–21, 1993.

[2] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[3] J. Wang, "Is artificial intelligence capable of understanding? an analysis based on philosophical hermeneutics," *Cultures of Science*, vol. 4, no. 3, pp. 135–146, 2021.

[4] S. Akhai, "From black boxes to transparent machines: The quest for explainable ai (ssrn scholarly paper 4390887)," 2023.

[5] A. Zador, S. Escola, B. Richards, B. Ölveczky, Y. Bengio, K. Boahen, M. Botvinick, D. Chklovskii, A. Churchland, C. Clopath, J. DiCarlo, S. Ganguli, J. Hawkins, K. Körding, A. Koulakov, Y. LeCun, T. Lillicrap, A. Marblestone, B. Olshausen, others, and D. Tsao, "Catalyzing next-generation artificial intelligence through neuroai," *Nature Communications*, vol. 14, no. 1, 2023, article 1.

[6] S. M. Kosslyn, W. L. Thompson, and G. Ganis, *The case for mental imagery (pp. vi, 248)*. Oxford University Press, 2006.

[7] A. M. Borghi, L. Barca, F. Binkofski, and L. Tummolini, "Varieties of abstract concepts: Development, use and representation in the brain," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 373, no. 1752, p. 20170121, 2018.

[8] D. W. Zaidel, "Creativity, brain, and art: Biological and neurological considerations," *Frontiers in Human Neuroscience*, vol. 8, 2014. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnhum.2014.00389

[9] K. Bobrowicz and J.-P. Thibaut, "The development of flexible problem solving: An integrative approach," *Journal of Intelligence*, vol. 11, no. 6, 2023, article 6.

[10] V. Balasubramanian, "Brain power," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 32, p. e2107022118, 2021, barlow, H. B. (2012). Possible Principles Underlying the Transformations of Sensory Messages. In W. A. Rosenblith (Ed.), Sensory Communication (pp. 216–234). The MIT Press. DOI: 10.7551/mitpress/9780262518420.003.0013.

[11] G. M. Reséndiz-Benhumea, E. Sangati, F. Sangati, S. Keshmiri, and T. Froese, "Shrunken social brains? a minimal model of the role of social interaction in neural complexity," *Frontiers in Neurorobotics*, vol. 15, p. 634085, 2021.

[12] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Inc: Henry Holt and Co, 1982.

[13] T. Parr, G. Pezzulo, and K. J. Friston, *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. The MIT Press, 2022.

[14] K. Roy, A. Jaiswal, and P. Panda, "Towards spike-based machine intelligence with neuromorphic computing," *Nature*, vol. 575, no. 7784, 2019, article 7784.

[15] W. Ou, S. Xiao, C. Zhu, W. Han, and Q. Zhang, "An overview of brain-like computing: Architecture, applications, and future trends," *Frontiers in Neurorobotics*, vol. 16, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnbot.2022.1041108

[16] G. Tang, A. Shah, and K. P. Michmizos, "Spiking neural network on neuromorphic hardware for energy-efficient unidimensional slam (arxiv:1903.02504). arxiv," 2019.

[17] C. Mayr, S. Hoeppner, and S. Furber, "Spinnaker 2: A 10 million core processor system for brain simulation and machine learning (arxiv:1911.02385). arxiv," 2019.

[18] C. Pehle, S. Billaudelle, B. Cramer, J. Kaiser, K. Schreiber, Y. Stradmann, J. Weis, A. Leibfried, E. Müller, and J. Schemmel, "The brainscales-2 accelerated neuromorphic system with hybrid plasticity," *Frontiers in Neuroscience*, vol. 16, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2022.795876

[19] C. Bartolozzi, G. Indiveri, and E. Donati, "Embodied neuromorphic intelligence," *Nature Communications*, vol. 13, no. 1, 2022, article 1.

[20] J.-J. Shu, Z. H. Tan, Q.-W. Wang, and K.-Y. Yong, "Programmable biomolecule-mediated processors," *Journal of the American Chemical Society*, vol. 145, no. 46, pp. 25 033–25 042, 2023.

[21] Y. Huo, J. Hu, Y. Yin, P. Liu, K. Cai, and W. Ji, "Self-assembling peptide-based functional biomaterials," *ChemBioChem*, vol. 24, no. 2, p. e202200582, 2023.

[22] L. Smirnova, B. S. Caffo, D. H. Gracias, Q. Huang, I. E. Morales Pantoja, B. Tang, D. J. Zack, C. A. Berlinicke, J. L. Boyd, T. D. Harris *et al.*, "Organoid intelligence (oi): the new frontier in biocomputing and intelligence-in-a-dish," *Frontiers in Science*, vol. 1, p. 1017235, 2023.

[23] P.-Y. Oudeyer, "Autonomous development and learning in artificial intelligence and robotics: Scaling up deep learning to human-like learning," *Behavioral and Brain Sciences*, vol. 40, p. e275, 2017.

[24] D. Vernon, "Cognitive vision: The case for embodied perception," *Image and Vision Computing*, vol. 26, no. 1, pp. 127–140, 2008.

[25] A. Cangelosi and T. Riga, "An embodied model for sensorimotor grounding and grounding transfer: Experiments with epigenetic robots," *Cognitive Science*, vol. 30, no. 4, pp. 673–689, 2006.

[26] A. Cangelosi and M. Schlesinger, *Developmental Robotics: From Babies to Robots*. MIT Press, 2015.

[27] L. Pecyna, A. Cangelosi, and A. Di Nuovo, "A robot that counts like a child: A developmental model of counting and pointing," *Psychological Research*, vol. 86, no. 8, pp. 2495–2511, 2022.

[28] D. Proudfoot, "Child machines," in *The Turing Guide (p. 0)*, J. Copeland, J. Bowen, M. Sprevak, and R. Wilson, Eds. Oxford University Press, 2017.

[29] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, "Cognitive developmental robotics: A survey," *IEEE Transactions on Autonomous Mental Development*, vol. 1, no. 1, pp. 12–34, 2009.

[30] J. Piaget, *The Psychology of Intelligence*. Routledge, 2001.

[31] A. Di Nuovo and A. Cangelosi, "Abstract concept learning in cognitive robots," *Current Robotics Reports*, vol. 2, no. 1, pp. 1–8, 2021.

[32] L. S. Vygotsky, *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press, 1978.

[33] R. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal on Robotics and Automation*, vol. 2, no. 1, pp. 14–23, 1986.

[34] K. Friston, "The free-energy principle: A unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.

[35] O. Çatal, S. Wauthier, C. De Boom, T. Verbelen, and B. Dhoedt, "Learning generative state space models for active inference," *Frontiers in Computational Neuroscience*, vol. 14, 2020. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fncom.2020.574372

[36] R. Smith, P. Schwartenbeck, T. Parr, and K. J. Friston, "An active inference approach to modeling structure learning: Concept learning as an example case," *Frontiers in Computational Neuroscience*, vol. 14, 2020. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fncom.2020.00041

[37] N. Sajid, P. J. Ball, T. Parr, and K. J. Friston, "Active inference: Demystified and compared," *Neural Computation*, vol. 33, no. 3, pp. 674–712, 2021.

[38] P. Lanillos, C. Meo, C. Pezzato, A. A. Meera, M. Baioumy, W. Ohata, A. Tschantz, B. Millidge, M. Wisse, C. L. Buckley, and J. Tani, "Active inference in robotics and artificial agents: Survey and challenges (arxiv:2112.01871). arxiv," 2021.

[39] D. Kawahara, S. Ozeki, and I. Mizuuchi, "A curiosity algorithm for robots based on the free energy principle," in *2022 IEEE/SICE International Symposium on System Integration (SII)*, 2022, pp. 53–59.

[40] W. T. Powers, *Behavior: The control of perception*. Chicago: Aldine de Gruyter, 1973.

[41] A. Linson, A. Clark, S. Ramamoorthy, and K. Friston, "The active inference approach to ecological perception: General information dynamics for natural and artificial embodied cognition," *Frontiers in Robotics and AI*, vol. 5, 2018. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frobt.2018.00021

[42] L. Da Costa, P. Lanillos, N. Sajid, K. Friston, and S. Khan, "How active inference could help revolutionise robotics," *Entropy*, vol. 24, no. 3, p. 361, 2022.

[43] T. Parr and G. Pezzulo, "Understanding, explanation, and active inference," *Frontiers in Systems Neuroscience*, vol. 15, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnsys.2021.772641

[44] L. Pio-Lopez, A. Nizard, K. Friston, and G. Pezzulo, "Active inference and robot control: A case study," *Journal of the Royal Society, Interface*, vol. 13, no. 122, p. 20160616, 2016.

[45] C. Pezzato, R. Ferrari, and C. H. Corbato, "A novel adaptive controller for robot manipulators based on active inference," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2973–2980, 2020.

[46] T. Van de Maele, T. Verbelen, O. Çatal, C. De Boom, and B. Dhoedt, "Active vision for robot manipulators using the free energy principle," *Frontiers in Neurorobotics*, vol. 15, p. 642780, 2021.

[47] P. Lanillos and G. Cheng, "Adaptive robot body learning and estimation through predictive coding," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 4083–4090.

[48] G. Oliver, P. Lanillos, and G. Cheng, "An empirical study of active inference on a humanoid robot," *IEEE Transactions on Cognitive and Developmental Systems, PP*, pp. 1–1, 2021.

[49] O. Çatal, T. Verbelen, T. Van de Maele, B. Dhoedt, and A. Safron, "Robot navigation as hierarchical active inference," *Neural Networks*, vol. 142, pp. 192–204, 2021.

[50] D. Burghardt and P. Lanillos. (2021) Robot localization and navigation through predictive processing using lidar (arxiv:2109.04139). arxiv. [Online]. Available: http://arxiv.org/abs/2109.04139

[51] M. Baioumy, C. Pezzato, C. H. Corbato, N. Hawes, and R. Ferrari, "Towards stochastic fault-tolerant control using precision learning and active inference," in *Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, M. Kamp, I. Koprinska, A. Bibal, T. Bouadi, B. Frénay, L. Galárraga, J. Oramas, L. Adilova, Y. Krishnamurthy, B. Kang, C. Largeron, J. Lijffijt, T. Viard, P. Welke, M. Ruocco, E. Aune, C. Gallicchio, G. Schiele, and F. Pernkopf, Eds. Springer International Publishing, 2021, pp. 681–691.

[52] W. Ohata and J. Tani, "Investigation of the sense of agency in social cognition, based on frameworks of predictive coding and active inference: A simulation study on multimodal imitative interaction," *Frontiers in Neurorobotics*, vol. 14, p. 61, 2020.

[53] T. Isomura, H. Shimazaki, and K. J. Friston, "Canonical neural networks perform active inference," *Communications Biology*, vol. 5, no. 1, 2022, article 1.

[54] D. Gandolfi, F. M. Puglisi, G. M. Boiani, G. Pagnoni, K. J. Friston, E. D'Angelo, and J. Mapelli, "Emergence of associative learning in a neuromorphic inference network," *Journal of Neural Engineering*, vol. 19, no. 3, p. 036022, 2022.

[55] B. J. Kagan, A. C. Kitchen, N. T. Tran, F. Habibollahi, M. Khajehnejad, B. J. Parker, A. Bhat, B. Rollo, A. Razi, and K. J. Friston, "In vitro neurons learn and exhibit sentience when embodied in a simulated gameworld," *Neuron*, vol. 110, no. 23, pp. 3952–3969.e8, 2022.

[56] A. Benítez-Burraco, Z. Clay, and V. Kempe, "Editorial: Self-domestication and human evolution," *Frontiers in Psychology*, vol. 11, p. 2007, 2020.

[57] Q. Guo, P. Sun, M. Cai, X. Zhang, and K. Song, "Why are smarter individuals more prosocial? a study on the mediating roles of empathy and moral identity," *Intelligence*, vol. 75, pp. 1–8, 2019.

[58] A. Knafo-Noam, D. Vertsberger, and S. Israel, "Genetic and environmental contributions to children's prosocial behavior: Brief review and new evidence from a reanalysis of experimental twin data," *Current Opinion in Psychology*, vol. 20, pp. 60–65, 2018.

[59] M. Elinder and O. Erixson, "An inquiry into the relationship between intelligence and prosocial behavior: Evidence from swedish population registers (ssrn scholarly paper 4038027)," 2022.

[60] L. Barrett, S. P. Henzi, and R. A. Barton, "Experts in action: Why we need an embodied social brain hypothesis," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 377, no. 1844, p. 20200533, n.d.

[61] D. Kappel and C. Tetzlaff, "Synapses learn to utilize pre-synaptic noise for the prediction of postsynaptic dynamics," 04.22.489175). bioRxiv, Tech. Rep., 2022, p. 2022.

[62] K. Friston, T. FitzGerald, F. Rigoli, P. Schwartenbeck, and G. Pezzulo, "Active inference: A process theory," *Neural Computation*, vol. 29, no. 1, pp. 1–49, 2017.

[63] J. Kwisthout and I. van Rooij, "Computational resource demands of a predictive Bayesian brain," *Computational Brain & Behavior*, vol. 3, no. 2, pp. 174–188, 2020.

[64] S. P. Wilson, "Modelling the emergence of rodent filial huddling from physiological huddling," *Royal Society Open Science*, vol. 4, no. 11, p. 170885, 2017.

[65] C. Heins, B. Millidge, D. Demekas, B. Klein, K. Friston, I. Couzin, and A. Tschantz, "pymdp: A python library for active inference in discrete state spaces. arxiv," 2022.

[66] J. P. Burkett, E. Andari, Z. V. Johnson, D. C. Curry, F. B. de Waal, and L. J. Young, "Oxytocin-dependent consolation behavior in rodents," *Science*, vol. 351, no. 6271, pp. 375–378, 2016.

[67] R. Smith, K. J. Friston, and C. J. Whyte, "A step-by-step tutorial on active inference and its application to empirical data," *Journal of mathematical psychology*, vol. 107, p. 102632, 2022.

[68]

[69] anni.jakobsson@csc.fi, "Lumi ranked third on top500 list of world's fastest supercomputers – the fastest supercomputer in europe," *LUMI*, May 2022. [Online]. Available: https://www.lumi-supercomputer.eu/lumi-ranked-third-on-top500-list-of-worlds-fastest-supercomputers-the-fastest-superco

[70] L. Beckes, H. IJzerman, and M. Tops, "Toward a radically embodied neuroscience of attachment and relationships," *Frontiers in Human Neuroscience*, vol. 9, 2015. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnhum.2015.00266

[71] J. Bowlby, *A secure base: Parent-child attachment and healthy human development*. Basic Books, 1988.

[72] ——, *Attachment and loss. 1: Attachment (2*. ed): Basic Books, 2003.

[73] Y. Ding, X. Xu, Z. Wang, H. Li, and W. Wang, "The relation of infant attachment to attachment and cognitive and behavioural outcomes in early childhood," *Early Human Development*, vol. 90, no. 9, pp. 459–464, 2014.

[74] R. I. M. Dunbar and S. Shultz, "Evolution in the social brain. science," *New York, N. Y.)*, vol. 317, no. 5843, pp. 1344–1347, 2007.

[75] B. Hare and M. Tomasello, "Human-like social skills in dogs?" *Trends in Cognitive Sciences*, vol. 9, no. 9, pp. 439–444, 2005.

[76] D. Cittern, T. Nolte, K. Friston, and A. Edalat, "Intrinsic and extrinsic motivators of attachment under active inference," *PLoS One*, vol. 13, no. 4, p. e0193955, April 2018, pMID: 29621266. PMCID: PMC5886414.

[77] T. J. Prescott and J. M. Robillard, "Are friends electric? the benefits and risks of human-robot relationships," *iScience*, vol. 24, no. 1, p. 101993, 2021.