Geometric Design Principles for Brains of Embodied Agents

Nihat   Ay

SANTA FE INSTITUTE

Geometric Design Principles for Brains of Embodied Agents

Nihat Ay

February 21, 2015

Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany
Department of Mathematics and Computer Science, Leipzig University, 04009 Leipzig, Germany
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Abstract

I propose a formal model of the sensorimotor loop and discuss its extrinsic embodiment constraints in relation to intrinsic degrees of freedom. These degrees form the basis of adaptation in terms of learning and should therefore be coupled with the embodiment constraints. Notions of sufficiency and embodied universal approximation allow us to formulate principles for such a coupling. This provides a geometric approach to the design of control architectures for embodied agents.

Keywords: sensorimotor loop, embodiment, information geometry, cheap design, universal approximation

Contents

1 Introduction

Within the last few decades, it has become clear that an understanding of intelligence and cognition has to take into account the fact that behaving agents are embodied and situated [11, 24]. At first sight this might appear like an obvious and unimportant observation. After all, it is obvious that agents *do* have bodies and that they *are* situated in some environment. However, it turns out that this observation has actually far-reaching implications. Any behaviour of an agent is mediated through its sensorimotor apparatus and its interactions with the environment. Therefore, the control of intelligent behaviour is inseparably intertwined with the agent's sensorimotor constraints.

Such a coupling should allow us to reveal design principles for brains and their implemented control mechanisms. In this regard, it has been pointed out that quite complex behavioural patterns do not necessarily require complex control, leading to the notions of morphological computation and the principle of cheap design [24, 30, 29].

In this article, I want to formally address the tight connection between the control architecture and the embodiment of an agent in terms of geometry, in particular information geometry [2, 4]. In order to do so, I apply a theory of the sensorimotor loop in terms of a causal model, as developed in [15, 8, 18], and propose an approach for designing controller architectures that utilise the sensorimotor constraints. My aim is not to provide the most general results, but to exemplify the strength of the theory by integrating initial results from an existing body of work [3, 5, 9, 14, 16, 17, 18, 30], and discussing them, for the first time, from a unifying perspective. The main intuition that underlies the connection between control architectures and sensorimotor constraints is based on two assumptions. On one hand, any architecture should be sufficient for the generation of desired behaviours given the embodiment constraints. This clearly requires some amount of richness or complexity of the architecture. On the other hand, it should be concise in some natural sense. I discuss various versions of these two assumptions and derive corresponding results. In particular, I address the notion of embodied universal approximation and contrast it with the standard notion of universal approximation. This is closely related to the recent works [19, 18].

Section 2 provides the conceptual and formal definition of the sensorimotor loop of an embodied agent. To the agent, the world appears as a black box which contains the agent's body and its environment. Sensorimotor mechanisms and their generation of behaviour, which takes place in the world, are formalised. Section 3 discusses and formalises extrinsic and intrinsic sensorimotor constraints. Intrinsic constraints are given in terms of a controller model, which allows the agent to adapt to the extrinsic constraints. The notions of sufficiency and embodied universal approximation are introduced as basis for a geometric design of controller models. Finally, Section 4 exemplifies the developed geometric methods in the context of policy models and derives four examples of embodied universal approximators.

2 A Formal Model of Embodied Agents

2.1 The Basic Components of the Sensorimotor Loop

What is an embodied agent? In order to develop a theory of embodied agents that allows us to cast the core principles of the field of embodied intelligence into rigorous theoretical and quantitative statements, we need an appropriate formal model. Such a model should be general enough to be applicable to all kinds of embodied agents, including natural as well as artificial ones, and specific enough to capture the essential aspects of embodiment. How should such a model look like? First of all, obviously, an embodied agent has a body. This body is situated in an environment with which the agent can interact, thereby generating some behavior. In order to be useful, this behavior has to be guided or controlled by the agent's brain or controller. We draw the boundary between the brain on one side and the body, together with the environment, on the other side. We assume that the brain receives sensor signals from and sends actuator signals to the outside world (body and environment), which are the only signal transmissions between the brain and the world. Stated differently, we assume that the brain is causally independent of the world, given the sensor signals, and the world is causally independent of the brain, given the actuator signals. We refer to this perspective as the *black box perspective*. In particular, the boundary between the body and the environment is not directly "visible" for the brain and has to be identified or actively constructed as result of the interaction with the world.
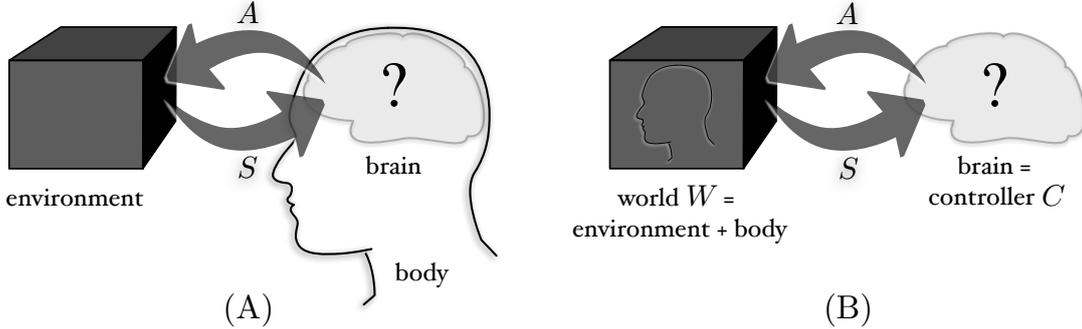
Figure 1: Two ways to define the agent separation: (A) agent (= brain + body) interacts with its environment considered as a black box; (B) agent (= brain) interacts with the world (= environment + body) considered as a black box.

2.2 The Mechanisms of the Sensorimotor Loop

Let us now develop a formal description of this sensorimotor loop. We denote the set of world states by $\mathscr{W}$. This set can be, for instance, the position of a robot in a static 3D environment. Information from the world is transmitted to the brain through sensors. Denoting the set of sensor states by $\mathscr{S}$, we model the mechanism of a sensor in terms of an information transmission channel from $\mathscr{W}$ to $\mathscr{S}$ as it is defined in information theory. More precisely, given a world state $w \in \mathscr{W}$, the response of the sensor can be characterized by a probability distribution of possible sensor states $s \in \mathscr{S}$ as result of w. For instance, if the sensor is noisy, then its response will not be uniquely determined. If the sensor is noiseless, that is deterministic, then there will be only one sensor state as possible response to the world state w. In any case, the response of the sensor given the world state w can be described as a distribution $\beta(w)$ on the sensor states s. Therefore, the mechanism of the sensor can be formalised as a Markov kernel

$$\beta : \ \mathscr{W} \ \longrightarrow \ \Delta_{\mathscr{S}} \,,$$

where $\Delta_{\mathscr{S}}$ denotes the set of distributions on $\mathscr{S}$. Each distribution $\beta(w)$ is uniquely determined by its values $\beta(w; ds)$ for infinitesimal sensor state sets ds. From this, we can calculate the probability of any (measurable) set $\mathscr{S}' \subseteq \mathscr{S}$ by integration:

$$\beta(w; \mathscr{S}') \ = \ \int_{\mathscr{S}'} \beta(w; ds) \,.$$

Given a reference measure μ_S on the set of sensor states so that all distributions $\beta(w)$ are continuous with respect to μ_S, we write $\beta(w; s)$ for the corresponding density of $\beta(w)$, that is $\beta(w; \mathscr{S}') = \int_{\mathscr{S}'} \beta(w; s) \, \mu_S(ds)$. In particular, this always holds in the case of finitely many sensor states where we take as reference measure the counting measure. Finally, we denote the set of all sensor mechanisms by $\Delta_{\mathscr{S}}^{\mathscr{W}}$.

After having described in detail the mathematical model of a sensor, it is now straightforward to consider corresponding formalizations of the other components of the sensorimotor loop in terms of Markov kernels, where we apply, without explicitly stating, the same notation conventions as for the sensor. Based on our conceptual perspective, which is illustrated in Fig. 1 (B), we compose the sensorimotor mechanisms according to a causal diagram, which represents the interaction process of the agent with the world (see Fig. 2).

We now continue with the notion of a policy. The agent can generate an effect in the world in terms of its actuators. Since we consider the body as being part of the world, this can lead, for instance, to some body movement of the agent. In order to control or guide this movement, it is beneficial for the agent to choose its actuator state based on the internal controller state which
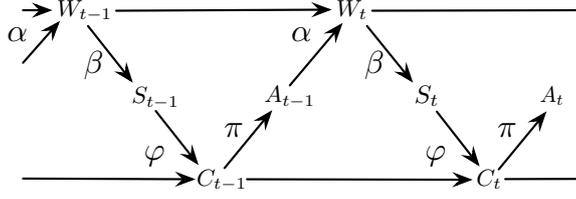
Figure 2: The causal diagram of the sensorimotor loop as conceptually illustrated in Fig. 1 (B).

contains information about the world received through its sensors. Denoting the controller state set by $\mathscr{C}$ and the actuator state set by $\mathscr{A}$, we can again consider a channel from $\mathscr{C}$ to $\mathscr{A}$ as formal model of a policy, which we denote by π. Being more precise, a policy receives a controller state c and generates, based on c, a distribution $\pi(c)$ of actuator states. Again, we have a Markov kernel

$$\pi : \ \mathscr{C} \ \longrightarrow \ \Delta_{\mathscr{A}} \, .$$

Note that this definition of a policy allows us to also consider a random choice of actions, so-called non-deterministic policies.

Finally, we formalise the dynamics α of the world and the dynamics φ of the controller, in terms of Markov kernels

$$\alpha : \ \mathscr{W} \times \mathscr{A} \ \longrightarrow \ \Delta_{\mathscr{W}} \qquad \text{and} \qquad \varphi : \ \mathscr{C} \times \mathscr{S} \ \longrightarrow \ \Delta_{\mathscr{C}} \, .$$

2.3 From Mechanisms to Embodied Behaviours

In order to describe the interaction process between the agent and the world, we have to sequentially apply the individual mechanisms in the right order. We first consider only one step as shown in Fig. 3, and then iterate this step. Starting with an initial world state w and an initial controller
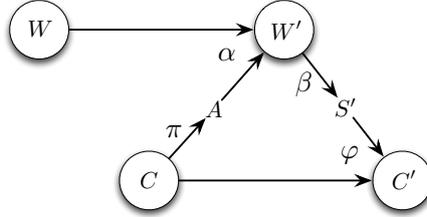
Figure 3: The one step mechanism of going from w, c to w', c' as defined by (1).

state c, we have the following Markov kernel that describes the transition to the new world and controller states:

$$
\begin{aligned}
\mathbb{P}(w, c; dw', dc') &= \int_{\mathscr{A}} \int_{\mathscr{S}} \varphi(c, s'; dc') \, \beta(w'; ds') \, \alpha(w, a'; dw') \, \pi(c; da') \\
&= \int_{\mathscr{S}} \varphi(c, s'; dc') \, \beta(w'; ds') \int_{\mathscr{A}} \alpha(w, a'; dw') \, \pi(c; da') \\
&= \mathbb{P}(c, w'; dc') \, \mathbb{P}(w, c; dw') \, .
\end{aligned}
\tag{1}
$$

Iterating this one-step transition T times defines a joint process $(w^1, c^1), \ldots, (w^T, c^T)$ of the world and the controller, conditioned on the initial joint state (w^0, c^0). Behaviour is a process that takes place in the world, for instance as a particular movement of the agent's body which we considered to be part of the world (see Fig. 1). This implies that only the world process $w^1, \ldots, w^T$ will be

4

of relevance for the study of behaviour. Marginalising out the controller process $c^1, \ldots, c^T$ leads to

$$
\begin{aligned}
\mathbb{P}(w^0, c^0; dw^1, \ldots, dw^T) &= \int_{\mathscr{C}} \cdots \int_{\mathscr{C}} \mathbb{P}(w^0, c^0; dw^1, dc^1, \ldots, dw^T, dc^T) \\
&= \int_{\mathscr{C}} \cdots \int_{\mathscr{C}} \prod_{t=1}^{T} \mathbb{P}(w^{t-1}, c^{t-1}; dw^t, dc^t),
\end{aligned}
\tag{2}
$$

where we integrate T times. This Markov kernel encodes all information that is required for the evaluation of behaviour.

3 Intrinsic and Extrinsic Sensorimotor Constraints

In the previous section, I have introduced the sensorimotor loop in terms of the mechanisms π, α, β, and φ. Here, the mechanisms α and β are extrinsic to the agent and define its embodiment constraints. I refer to these constraints as being of the first kind in order to distinguish them from other constraints introduced below. Clearly, α and β are not fixed and can evolve in time. For instance, if an agent moves from one environment to another the mechanism α will change. Both mechanisms, α and β, will change, if the agent's body is subject to a developmental process or to a partially disabling injury. However, we can expect that, due to invariant structures and regularities of the environment and the body, the mechanisms will remain in a restricted domain. This fact implies a second kind of embodiment constraint. In addition to the constraints of the first kind encoded by a single pair $\sigma := (\alpha, \beta)$ of extrinsic mechanisms, the evolution of σ is restricted to trajectories within a subset Σ of possible mechanisms, corresponding to the variability of the body and the environment. We consider the direct cause of this evolution within Σ to be completely exogenous with respect to the agent and, therefore, it can not be directly controlled by the agent. In contrast, we assume that the intrinsic mechanisms π and φ, summarised as $\gamma = (\pi, \varphi)$, can be endogenously modified by the agent in terms of a learning process. The question is whether this intrinsic process is also constrained. Obviously, this will depend on the degrees of freedom of the controller or brain architecture. If it is rich enough to implement any π and φ then we have to consider the set

$$
\Sigma \times \Gamma, \qquad \text{where } \Gamma := \Delta_{\mathscr{A}}^{\mathscr{C}} \times \Delta_{\mathscr{C}}^{\mathscr{C} \times \mathscr{S}},
\tag{3}
$$

as domain for the process of all sensorimotor mechanisms. But should we assume this richness of brain architectures? I argue that it is not only not required but also beneficial for the agent to have a restricted control architecture, which is in line with the idea of cheap design within the field of embodied intelligence. However, the theoretical study of unrestricted control architectures, referred to as *universal approximators*, is helpful for exploring the ones that are optimal for the given constraints Σ of the agent. Here, optimality involves two requirements. On one hand, the control architecture should be sufficient in the sense that it enables the agent to adapt to the embodiment constraints Σ. On the other hand, it should be concise in order to efficiently implement this adaptation. The field of embodied intelligence offers many case studies as evidence for such kinds of cheap control. This field highlights, in particular, the fact that quite complex and useful behaviours do not require much control [24].

Now, I am going to make these ideas on optimal architectures more precise. A brain architecture can be associated with a set $\mathcal{M} \subseteq \Gamma$, the set of those $\gamma = (\pi, \varphi)$ that can be generated by the architecture. We refer to such a set as *controller model* or simply *model*. Most generally, a model can be any subset $\mathcal{M}$ of Γ. Typically, a model $\mathcal{M}$ is defined as the image of a map η, referred to as a *parametrisation* of $\mathcal{M}$, from a parameter set $\Theta \subseteq \mathbb{R}^d$ to Γ. Fig. 4 illustrates the important parametrisation in terms of synaptic couplings w_{ij} of neurons i and j. We call a model $\mathcal{M}$ together with a parametrisation η a *parametrised model*. Clearly, any map $\eta : \Theta \to \Gamma$ is a parametrisation of the model given by its image. In order to be useful in applications, a parametrised model is typically assumed to have further properties which are context dependent,
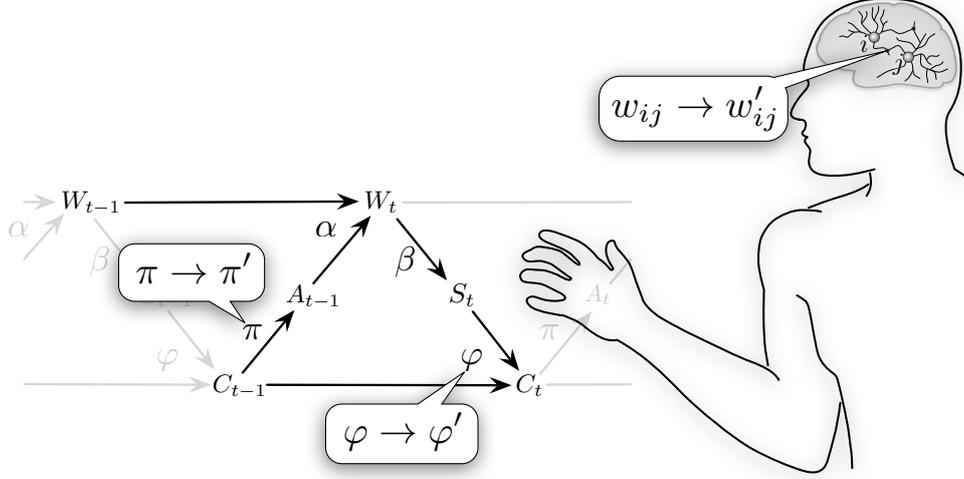
Figure 4: Changes of the intrinsic mechanisms π and φ as result of changes of parameters such as synaptic coupling strengths.

for instance smoothness properties up to some order.

In addition to the embodiment constraints Σ, a model $\mathcal{M}$ encodes the intrinsic constraints of the brain and represents, at the same time, the domain in which learning processes can take place. The set (3) of sensorimotor loops then reduces to

$$SML := \Sigma \times \mathcal{M}. \tag{4}$$

In what follows, I want to identify intrinsic constraints, that is a model $\mathcal{M}$, that are coupled with the extrinsic constraints Σ according to some natural optimality conditions. In order to do so, let us consider the map that assigns to each sensorimotor loop its corresponding behaviour, the *behaviour map*:

$$\psi_\sigma : \Gamma \longrightarrow \Delta_{\mathscr{W}\infty}^{\mathscr{W}\times\mathscr{C}}, \qquad \gamma \longmapsto \mathbb{P}(w^0, c^0; dw^1, dw^2, dw^3, \dots), \tag{5}$$

where we use the formula (2) to define the Markov kernel $\psi_\sigma(\gamma)$ by letting T go to ∞. Now, say that we have given for all $\sigma \in \Sigma$ a set $\mathcal{O}_\sigma$ of optimal or desired behaviours, such as behaviours with maximal predictive information [30] or maximal expected reward [9]. In this article, optimality of behaviours is not further specified and should not be confused with the optimality of a model, which plays the central role in this paper. Optimal models should, at least, satisfy the following natural sufficiency condition.

Definition 1. We say that a model $\mathcal{M}$ is (*geometrically*) *sufficient*, if for all $\sigma \in \Sigma$ and all corresponding behaviours $\delta \in \mathcal{O}_\sigma$, there exists $\gamma \in \overline{\mathcal{M}}$ that generates that behaviour, that is $\psi_\sigma(\gamma) = \delta$.

Here, the bar over the set $\mathcal{M}$ denotes its topological closure in Γ. In principle, this will depend on the underlying topology of Γ for which one has various natural choices. I am not going to address these topological questions in further detail. The main results of the next section refer to the case where the state sets are finite and therefore the topology is simply the standard one of a finite-dimensional real vector space. If the closure of the model $\mathcal{M}$ equals all theoretically possible intrinsic mechanisms, that is $\overline{\mathcal{M}} = \Gamma$, then we say that the model is a *universal approximator*. This corresponds to the most flexible brain architecture which I already mentioned above

6

(see equation (3)). I argue that such a brain is not required for embodied agents in order to be universal at the behavioural level. There are sufficient models, in the sense of Definition 1, that are behaviourally equivalent to, but less complex than, a universal approximator. Clearly one has to specify the notion of complexity here. In any case, the general study of sufficient models in relation to their complexity provides one way to formally address the subject of cheap design which plays a central role within the field of embodied intelligence [24].

In order to be more precise, we consider as optimal behaviour sets $\mathcal{O}_\sigma$ all behaviours that can in principle be generated in the context of the extrinsic constraints given by σ. Stated differently, we set $\mathcal{O}_\sigma$ to be the image of ψ_σ and obtain the following sufficiency property of a model $\mathcal{M}$, which we then call an *embodied universal approximator (for Σ)*:

$$\psi_\sigma(\overline{\mathcal{M}}) = \psi_\sigma(\Gamma) \qquad \text{for all } \sigma \in \Sigma \,. \tag{6}$$

In general, an embodied universal approximator does not have to be concise. For instance, any universal approximator is in particular an embodied universal approximator. In order to be consitent with the idea of cheap design, we have to specify what we mean by "cheap" in terms of a complexity measure and then compare various embodied universal approximators with respect to their complexity. This will be exemplified in the next section.

4 Cheap Embodied Universal Approximation

The intention of this section is to exemplify the developed concepts and tools in terms of very specific theoretical results on cheap design which are closely related to the work [18]. In particular, I assume only finite sets $\mathcal{W}$, $\mathcal{S}$, $\mathcal{C}$, and $\mathcal{A}$, and, furthermore, I concentrate on the definition of policy models, ignoring other intrinsic constraints. To be more precise, consider the set

$$\Gamma = \Delta_{\mathcal{A}}^{\mathcal{C}} \times \Delta_{\mathcal{C}}^{\mathcal{C} \times \mathcal{S}}$$

of all possible intrinsic mechanisms π and φ (see equation (3)). In general, a model $\mathcal{M}$ defines a reduction of Γ in both components, that is $\mathcal{M} \subseteq \Gamma$. In this section, however, we study only reductions in the first component in terms of policy models $\mathcal{M}_{\text{policy}} \subseteq \Delta_{\mathcal{A}}^{\mathcal{C}}$ which correspond to joint controller models

$$\mathcal{M} = \mathcal{M}_{\text{policy}} \times \Delta_{\mathcal{C}}^{\mathcal{C} \times \mathcal{S}} \,. \tag{7}$$

We say that a policy model is an *embodied universal approximator (for Σ)*, if the corresponding joint model (7) has this property. Note that the restriction to policy models excludes the possibility of coupling the two intrinsic mechanisms π and φ. Such a coupling might provide a further way of reducing the complexity of the controller model.

4.1 General Selection of Policy Models

Consider the one-step kernel $\mathbb{P}(w, c; dw')$, which appears in (1). For finite sets, this reduces to

$$\mathbb{P}^\pi(w, c; w') = \sum_a \alpha(w, a; w') \, \pi(c; a) \,, \qquad c \in \mathcal{C}, \ w, w' \in \mathcal{W} \,, \tag{8}$$

where we indicate the dependence on π explicitly. We leave all other mechanisms of the sensori-motor loop fixed and consider two policies π_1 and π_2 that satisfy

$$\mathbb{P}^{\pi_1}(w, c; w') = \mathbb{P}^{\pi_2}(w, c; w') \ \text{ for all } c \in \mathcal{C} \text{ and all } w, w' \in \mathcal{W} \,. \tag{9}$$

It follows directly from (1) that the corresponding behaviours, defined by (5), are identical. In other words, if for all controller states c the distributions $\pi_1(c; \cdot)$ and $\pi_2(c; \cdot)$ give the same expectation values of the functions

$$\alpha_{w,w'} : \ a \ \mapsto \ \alpha_{w,w'}(a) := \alpha(w, a; w') \,, \qquad w, w' \in \mathcal{W}, \tag{10}$$

then π_1 and π_2 will generate the same behaviour, assuming that all the other mechanisms are fixed. This redundancy allows us to construct policy models with smaller dimensionality without reducing the behavioural repertoire of the agent. In order to do so, we identify two distributions μ_1 and μ_2 on $\mathscr{A}$ if they give the same expectation values of the functions $\alpha_{w,w'}$, $w, w' \in \mathscr{W}$, and thereby obtain a partition of $\Delta_{\mathscr{A}}$ into convex equivalence classes $[\mu]$, $\mu \in \Delta_{\mathscr{A}}$. Given a policy π, we can now consider the family of classes $[\pi(c;\cdot)]$, $c \in \mathscr{C}$, and any further π' satisfying $\pi'(c;\cdot) \in [\pi(c;\cdot)]$ will generate the same behaviour as π. Following the idea of cheap design, it is natural to define a model by choosing π' to be the one with the least complexity. Assuming that unstructured low complexity policies have high entropy, we have the natural choice provided by the maximum entropy principle. More precisely, in each class $[\pi(c;\cdot)]$ we choose the distribution with the highest entropy $-\sum_{a \in \mathscr{A}} \pi(c;a) \ln \pi(c;a)$. This policy is uniquely determined, as the entropy is strictly concave and the classes are convex. We obtain our first parametrisation $\eta : \mathbb{R}^{\mathscr{C} \times \mathscr{W}^2} \to \Delta_{\mathscr{A}}^{\mathscr{C}}$,

$$\theta \; \mapsto \; \pi_\theta(c;a) := \frac{\exp\left(\sum_{w,w'} \theta_{w,w'}(c)\,\alpha_{w,w'}(a)\right)}{\sum_{a'} \exp\left(\sum_{w,w'} \theta_{w,w'}(c)\,\alpha_{w,w'}(a')\right)} \,. \tag{11}$$

This is a quite prominent model in statistics and information geometry, which is called *exponential family* [2]. Using general information-geometric arguments, it is obvious that the parametrisation (11) defines an embodied universal approximator for

$$\Sigma_\alpha \; := \; \left\{ (\alpha, \beta) \; : \; \beta \in \Delta_{\mathscr{S}}^{\mathscr{W}} \right\} \,. \tag{12}$$

On the other hand, in most interesting cases the number $|\mathscr{C}|\,|\mathscr{W}|^2$ of parameters of this model will be a huge. It turns out, however, that this "full" representation of policies is highly redundant. One obvious redundancy follows from the fact that the $\alpha(w, a; w')$ are probability distributions in w', that is $\sum_{w'} \alpha_{w,w'}(a) = 1$ for all $w \in \mathscr{W}$ and all $a \in \mathscr{A}$. As the space of constant functions on $\mathscr{A}$ cancels out for each $c \in \mathscr{C}$, the number of parameters can be reduced to $|\mathscr{C}|\,(|\mathscr{W}|^2 - 1)$. I argue that a further, conceptually deeper, reduction can be expected in embodied systems. Formally, it is reflected by the fact that the sum

$$\sum_{w,w'} \theta_{w,w'}(c)\,\alpha_{w,w'}(a) \,, \tag{13}$$

which appears in (11), can be strongly simplified due to linear dependence of the vectors $\alpha_{w,w'} \in \mathbb{R}^{\mathscr{A}}$, $w, w' \in \mathscr{W}$. Let us consider two obvious reasons for this.

1. Assume that the world state w' can not be reached from the world state w in one step, that is

$$\alpha(w, a; w') = 0 \quad \text{for all } a \in \mathscr{A} \,. \tag{14}$$

   In this case, the corresponding term disappears from the sum (13). In the situation of an embodied agent, the majority of pairs (w, w') has this property, because the physical constraints of the sensorimotor loop exclude most of the transitions from w to w' in one step.

2. It is not necessary that (14) holds in order to ignore a term from the sum (13). It is already sufficient that there is a constant $r \in \mathbb{R}$ such that

$$\alpha(w, a; w') = r \quad \text{for all } a \in \mathscr{A} \,. \tag{15}$$

   Although formally this property is a simple extension of (14), it highlights another important aspect. If $\alpha(w, a; w') = r > 0$ for all $a \in \mathscr{A}$ then w' can be reached from w in one step but in a way that does not involve the actuators. The transition from w to w' is not sensitive to the actuators, and therefore its representation does not play any role in our parametrisation (11).

8

Clearly, we can adjust the parametrisation (11) to the dimension d_α of the vector space

$$V_\alpha := \mathrm{span}\{\alpha_{w,w'} : w, w' \in \mathscr{W}\} \subseteq \mathbb{R}^{\mathscr{A}},$$

which is at most $|\mathscr{A}|$. Furthermore, note that the constant functions are contained in V_α. This implies that there are $d_\alpha - 1$ functions $\alpha_k \in V_\alpha$, which we refer to as *feature vectors*, such that every policy of the structure (11) can be expressed in terms of a linear combination of the α_k and a constant function. This leads to the following simplification of the parametrisation (11).

Proposition 2. *The parametrisation* $\eta : \mathbb{R}^{\mathscr{C}} \times \mathbb{R}^{d_\alpha - 1} \to \Delta_{\mathscr{A}}^{\mathscr{C}}$,

$$\theta \mapsto \pi_\theta(c; a) := \frac{\exp\left(\sum_{k=1}^{d_\alpha - 1} \theta_k(c) \, \alpha_k(a)\right)}{\sum_{a'} \exp\left(\sum_{k=1}^{d_\alpha - 1} \theta_k(c) \, \alpha_k(a')\right)}, \tag{16}$$

defines an embodied universal approximator for Σ_α.

In this parametrisation of the policies, we have at most $|\mathscr{C}|(d_\alpha - 1)$ parameters which can be much smaller than the number of parameters in the parametrisation (11). However, clearly they define the same policy model, which is a maximum entropy model. This was motivated by the idea that the maximum entropy distributions correspond to the least structured ones. However, high entropy distributions tend to have a large support. We can take a different standpoint and ask the following question: If an agent can generate a behaviour with only a few actuator states, why should it involve more actuator sates than that? In other words, if we consider actuator states to be costly then we should design the model in a way that the policies have only a minimal required number of actuator states. Interpreting, for simplicity of the argument, entropy as a cardinality measure, then we would rather prefer distributions with minimal entropy. From this perspective, the minimum entropy distributions appear more natural. Note, that, in contrast to the uniqueness of the maximum entropy distribution in a class $[\mu]$ where $\mu \in \Delta_{\mathscr{A}}$, there are many distributions that locally minimise the entropy in that class. On the other hand, it turns out that the cardinality of the support of all these distributions is upper bounded by d_α (see corresponding derivations in [3]). More precisely, the following holds: If for a policy π, all distributions $\pi(c; \cdot)$ are local minimisers of the entropy, then

$$\left|\{a \in \mathscr{A} : \pi(c; a) > 0\}\right| \leq d_\alpha \quad \text{for all } c \in \mathscr{C}. \tag{17}$$

Clearly, a policy model $\mathcal{M}_{\mathrm{policy}}$ is an embodied universal approximator, if it contains all policies that satisfy (17) in its closure. We now explicitly define such a policy model with

$$\dim \mathcal{M}_{\mathrm{policy}} \leq 2 |\mathscr{C}| d_\alpha.$$

Proposition 3. *For any embedding* $f : \mathscr{A} \to \mathbb{R}$, *that is* $f(a) = f(a')$ *only if* $a = a'$, *the parametrisation* $\eta : \mathbb{R}^{\mathscr{C}} \times \mathbb{R}^{2d_\alpha} \to \Delta_{\mathscr{A}}^{\mathscr{C}}$,

$$\theta \mapsto \pi_\theta(c; a) := \frac{\exp\left(\sum_{k=1}^{2d_\alpha} \theta_k(c) \, f^k(a)\right)}{\sum_{a'} \exp\left(\sum_{k=1}^{2d_\alpha} \theta_k(c) \, f^k(a')\right)}, \tag{18}$$

defines an embodied universal approximator for Σ_α.

The proof of this statement is based on [3, 16] and is given in the Appendix. The policy model defined by (18) has $2|\mathscr{C}| d_\alpha$ parameters, which is larger than the number $|\mathscr{C}|(d_\alpha - 1)$ of parameters of the previous model defined in Proposition 2. However, there are important differences. The feature vectors f^k, $k = 1, \ldots, 2d_\alpha$, do not explicitly depend on α, but their number depends on the dimension of V_α. Therefore, they can be used for all mechanisms α' that have at most the

same dimension $d_{\alpha'}$ as α. Furthermore, these feature vectors do not require much information to be specified. More precisely, the property of f that is required in Proposition 3 is a generic property of real functions on $\mathscr{A}$ so that f can be chosen randomly. With the set

$$\Sigma_{d_\alpha} := \left\{ (\alpha', \beta) \; : \; d_{\alpha'} \le d_\alpha, \; \beta \in \Delta_{\mathscr{S}}^{\mathscr{W}} \right\} \tag{19}$$

of extrinsic constrains, which is larger than Σ_α as defined by (12), we have the following direct implication of Proposition 3.

Corollary 4. *The policy model of Proposition 3 is an embodied universal approximator for Σ_{d_α}.*

4.2 From Models to Architectures

In the last section, we defined policy models for general unstructured sets $\mathscr{C}$ and $\mathscr{A}$, and the policies were abstract objects without any reference to a mechanism that generates or computes an output $a \in \mathscr{A}$ given the input $c \in \mathscr{C}$. In this section, we extend our previous ideas to the setting of systems consisting of interacting nodes such as neuronal systems. However, we restrict attention to the actuators and assume, for simplicity, that we have n binary actuators with states -1 and $+1$, that is $\mathscr{A} = \{-1, +1\}^n$. For each non-empty subset $N \subseteq [n] := \{1, \ldots, n\}$, the interaction among the actuators in N is described in terms of the monomial

$$a = (a_1, \ldots, a_n) \; \mapsto \; \theta_N(c) \prod_{i \in N} a_i \,, \tag{20}$$

where $\theta_N(c)$ is the interaction coefficient of the actuators in N, which we assume to be dependent on the controller state $c \in \mathscr{C}$. We refer to the cardinality of N as the *order* of the interaction. For $|N| = 2$, we have the important special case of a pairwise interaction, which is of particular interest within the field of neural networks. There, the interaction coefficients are usually interpreted as the synaptic connection strengths between the neurons. If we want to incorporate interactions among nodes of all subsets $N \subseteq [n]$, then we have to consider the following sum of monomials (20):

$$\sum_{\emptyset \ne N \subseteq [n]} \theta_N(c) \prod_{i \in N} a_i \,. \tag{21}$$

Note that, for each $c \in \mathscr{C}$ we have $2^n - 1$ parameters $\theta_N(c)$, so that the overall number of parameters is $|\mathscr{C}| \, (|\mathscr{A}| - 1)$, which coincides with the dimension of the polytope $\Delta_{\mathscr{A}}^{\mathscr{C}}$. Clearly, any function $\mathscr{A} = \{-1, +1\}^n \to \mathbb{R}$ can be written in the form (21). If we set in (21) all $\theta_N(c)$ to zero whenever $|N| > k$, we reduce this space of functions to the space of functions that involve at most interactions of order k. This defines the following policy model, which we refer to as *k-interaction model*:

$$\theta \; \mapsto \; \pi_\theta(c; a) = \frac{\exp\left(\sum_{\substack{\emptyset \ne N \subseteq [n] \\ |N| \le k}} \theta_N(c) \prod_{i \in N} a_i \right)}{\sum_{a'} \exp\left(\sum_{\substack{\emptyset \ne N \subseteq [n] \\ |N| \le k}} \theta_N(c) \prod_{i \in N} a_i' \right)} \,. \tag{22}$$

Given this increasing family of models, ordered according to k, we now address the problem of finding the sufficient order of interaction so that the corresponding k-interaction model is an embodied universal approximator for Σ_{d_α}. For a first simple estimate, we consider again the policy model of Proposition 3. Given that we now assume $\mathscr{A} = \{-1, +1\}^n$, it is possible to define the function f as a linear function, that is $f(a) = f(a_1, \ldots, a_n) = \sum_{i=1}^n w_i a_i$ so that $f(a) \ne f(a')$ whenever $a \ne a'$. For the k-th powers of f we obtain

$$f^k(a_1, \ldots, a_n) = \left(\sum_{i=1}^n w_i a_i \right)^k = \sum_{i_1, \ldots, i_k} w_{i_1} \cdots w_{i_k} a_{i_1} \cdots a_{i_k} \,.$$

This implies

$$\sum_{k=1}^{2\,d_\alpha} \theta_k(c)\, f^k(a_1,\dots,a_n) = \sum_{k=1}^{2\,d_\alpha} \theta_k(c) \sum_{i_1,\dots,i_k} w_{i_1} \cdots w_{i_k}\, a_{i_1} \cdots a_{i_k} = \sum_{\substack{N \subseteq [n] \\ |N| \le 2d_\alpha}} \theta_N(c) \prod_{i \in N} a_i$$

with corresponding coefficients $\theta_N(c)$. This proves that the policy model of Proposition 3 is contained in the $(2\,d_\alpha)$-interaction model, and that therefore the latter is also an embodied universal approximator. However, it turns out that this can be considerably improved. It directly follows from a result by Kahle [14] that $2\,d_\alpha$ can be replaced by a much smaller interaction order, which defines the following policy model.

Proposition 5. *With*

$$k(\alpha) := \lceil \log_2(d_\alpha + 1) \rceil,$$

the $k(\alpha)$-interaction model is an embodied universal approximator for Σ_{d_α} ($\lceil x \rceil$ denotes the smallest integer $\ge x$).

In neural networks theory, the interaction order is generally considered to be pairwise. This is valid for the policy model of Proposition 5 only in the trivial case $d_\alpha = 1$. If we aim at a neuronal interpretation of policy models then we have to represent higher order interactions in terms of pairwise interactions, which is always possible by extending the system appropriately. One important special case in line with this general idea is given by the so-called *restricted Boltzmann machine* (RBM). In order to define it, we extend the above system of n nodes by further m so-called *hidden nodes* with the same state space $\{-1, +1\}$. The overall state is then given as a pair $(h, a) = (h_1, \dots, h_m, a_1, \dots, a_n) \in \{-1, +1\}^{m+n}$. We now consider the following family of kernels,
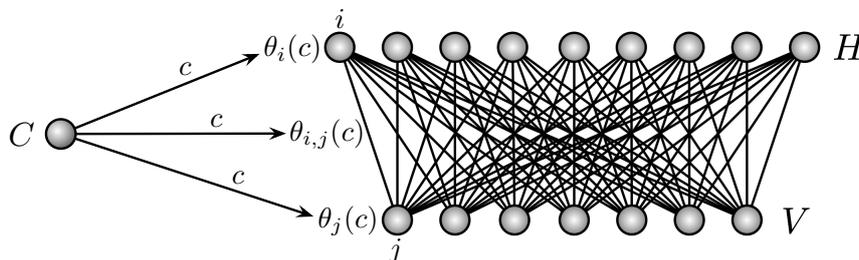

Figure 5: The architecture of a controlled restricted Boltzmann machine. It consists of a restricted Boltzmann machine with the set $V = \{1, \dots, n\}$ of visible nodes and the set $H = \{1, \dots, m\}$ of hidden nodes. The interaction is pairwise, where each node in V interacts with each node in H. No lateral interactions within V and within H are assumed. The interaction coefficients $\theta_{i,j}$, θ_i, and θ_j, where $i \in H$ and $j \in V$, are all functions of the controller state c.

which involves at most pairwise interactions between the hidden nodes and the actuators:

$$\pi_\theta(c; h, a) := \frac{\exp\left(\sum_{\substack{1 \le i \le m \\ 1 \le j \le n}} \theta_{i,j}(c)\, h_i\, a_j + \sum_{1 \le i \le m} \theta_i(c)\, h_i + \sum_{1 \le j \le n} \theta_j(c)\, a_j\right)}{\sum_{h',a'} \exp\left(\sum_{\substack{1 \le i \le m \\ 1 \le j \le n}} \theta_{i,j}(c)\, h'_i\, a'_j + \sum_{1 \le i \le m} \theta_i(c)\, h'_i + \sum_{1 \le j \le n} \theta_j(c)\, a'_j\right)}. \tag{23}$$

Finally, we marginalise out the hidden nodes and obtain the following policy model, which is, for each $c \in \mathscr{C}$, a restricted Boltzmann machine:

$$\theta \mapsto \pi_\theta(c; a) = \sum_h \pi_\theta(c; h, a). \tag{24}$$

Therefore, we call this policy model *controlled restricted Boltzmann machine* (see Fig. (5)). Due to general results by Le Roux & Bengio [27] and Montúfar & Ay [17], it is known that any distribution

11

of support cardinality smaller than or equal to d_α can be represented by an RBM with $d_\alpha + 1$ hidden nodes. This directly implies the following result.

Proposition 6. *A controlled restricted Boltzmann machine, defined by (23) and (24), with $m = d_\alpha + 1$ hidden nodes is an embodied universal approximator for Σ_{d_α}.*

The dependence of the parameters $\theta_{i,j}(c)$, $\theta_i(c)$, and $\theta_j(c)$ on the controller state c can be quite complicated. Assuming that the controller also has a composite structure with k binary nodes, it is possible to represent this dependence in terms of pairwise interactions between the controller nodes and the hidden nodes. This leads to the definition of *conditional restricted Boltzmann machines* [19]. Cheap control with such machines has been theoretically and experimentally studied in [18]. This study is based on the notion of an embodiment dimension, which is a refinement of the dimension d_α used in the present article.

5   Conclusions

In order to approach a general theory of embodied agents, I have introduced a formal model of the sensorimotor loop, which specifies its intrinsic and extrinsic mechanisms, building on previous work [15, 8, 18]. The extrinsic mechanisms represent the embodiment constraints of the system which can be utilised by appropriate adjustment of the intrinsic mechanisms in order to generate useful behaviours. This requires some degree of flexibility of the control architecture, which I formalised in terms of a sufficiency notion. As a particular case, I studied in more detail embodied universal approximation and the corresponding design of controller models in terms of geometry. However, I argued that sufficiency should not be the only requirement involved in systems design. In order to address the notion of cheap design within the field of embodied intelligence, we have to identify sufficient controller architectures with low complexity. Clearly, one would assume that selection pressure has generated such architectures of naturally evolved brains, which can cope with limitations of mass and energy resources. For instance, the field of sparse coding has provided evidence for internal representations of external stimuli in terms of sparse neuronal activity [23]. However, there are many aspects that contribute to the overall complexity of a controller model, and we are far from a conclusive definition of complexity that would account for the right notion of cheap control. Therefore, I was not very precise in this regard and used a few complexity notions as motivation for the policy models that I have defined. In this context, low complexity of a policy meant: 1. high entropy of the policy, 2. low number of actuator states used by the policy, 3. low interaction order among the actuators, and 4. low number of hidden nodes in a restricted Boltzmann machine.

There is a further important aspect of complexity, which I did not address explicitly. If we design policy models, we have to distinguish between the complexity of the model itself and the complexity of a particular policy, taken from that model. One can design models for which the individual policies have low complexity, but the overall model is hard to describe. However, it is problematic to implement such a structure. In nature, its information has to be transmitted through genetic inheritance. This information transmission clearly has a limited capacity, which is, for instance, modulated by the mutation rate. Therefore, robustness issues have also to be taken into account [6, 25]. To conclude, the right choice of a model should balance the complexities of both, the model itself and the policies that are implemented by the model. This is a quite natural idea within complexity theory [26, 12, 28, 13, 7].

Acknowledgements

was not used in the original proof [16], leading to an improvement of the original result. This work has been supported by the DFG Priority Program *Autonomous Learning*.

6 Appendix

Proof. **(Proposition 3)** We consider a policy π with the property that the supports $\mathscr{A}_c$ of its distributions $\pi(c;\cdot)$, $c \in \mathscr{C}$, on $\mathscr{A}$ have at most d_α elements. For each $c \in \mathscr{C}$, we define the function $g_c : \mathscr{A} \to \mathbb{R}$,

$$a \;\mapsto\; g_c(a) \;:=\; \prod_{a' \in \mathscr{A}_c} \left(f(a) - f(a')\right)^2 \;=\; \sum_{k=0}^{2\,d_\alpha} \theta_k^{(0)}(c)\, f^k(a)\,.$$

Obviously, $g_c \geq 0$, and $g_c(a) = 0$ if and only if $a \in \mathscr{A}_c$. Furthermore, we consider the polynomial $h_c : \mathbb{R} \to \mathbb{R}$, $t \mapsto \sum_{k=0}^{d_\alpha - 1} \theta_k^{(1)}(c)\, t^k$, that satisfies

$$\sum_{k=0}^{d_\alpha - 1} \theta_k^{(1)}(c)\, f^k(a) \;=\; h_c(f(a)) \;=\; \ln \pi(c;a), \qquad a \in \mathscr{A}_c\,.$$

Obviously, the one parameter family

$$\pi_\lambda(c;a) \;=\; \frac{e^{h_c(f(a)) - \lambda\, g_c(a)}}{\sum_{a' \in \mathscr{A}} e^{h_s(f(a')) - \lambda\, g_c(a')}}$$

of policies is contained in the model defined by (18). Note that we can ignore the term for $k = 0$ because this gives a constant which cancels out. Furthermore, it is easy to see that

$$\lim_{\lambda \to \infty} \pi_\lambda(c;a) \;=\; \lim_{\lambda \to \infty} \frac{e^{h_c(f(a))} \left(e^{-g_c(a)}\right)^\lambda}{\sum_{a' \in \mathscr{A}} e^{h_c(f(a'))} \left(e^{-g_c(a')}\right)^\lambda} \;=\; \pi(c;a)\,. \tag{25}$$

This proves that π is in the closure of the model defined by (18). $\qquad\square$

References

[1] S. Amari, Information geometry on hierarchy of probability distributions, IEEE Transactions on Information Theory 47 (2001) 1701–1711

[2] S. Amari, H. Nagaoka, Methods of Information Geometry, Oxford University Press 2000.

[3] N. Ay, An Information-Geometric Approach to a Theory of Pragmatic Structuring, The Annals of Probability 30 (1) (2002) 416–436

[4] N. Ay, J. Jost, H. V. Lê, L. Schwachhöfer, Information Geometry, Springer, submitted.

[5] N. Ay, A. Knauf, Maximizing Multi-Information, Kybernetika 42 (5) (2007) 517–538

[6] N. Ay, D.C. Krakauer, Geometric Robustness Theory and Biological Networks, Theory in Biosciences (2007) 93–121

[7] N. Ay, M. Müller, A. Szkoła, Effective complexity and its relation to logical depth, IEEE transactions on information theory 56 (9) (2010) 4593–4607

[8] N. Ay, K. Zahedi, On the Causal Structure of the Sensorimotor Loop, in: M. Prokopenko (ed.) Guided Self-Organization: Inception, Springer 2014.

[9] N. Ay, G. Montúfar, J. Rauh, Selection Criteria for Neuromanifolds of Stochastic Dynamics, Post-conference proceedings Advances in Cognitive Neurodynamics (III), Springer 2012

[10] N. Ay, E. Olbrich, N. Bertschinger, J. Jost, A Geometric Approach to Complexity, Chaos 21 (2011) 037103

[11] R. A. Brooks, Intelligence without representation, Artificial Intelligence 47 (1-3) (1991) 139–159

[12] M. Gell-Mann, S. Lloyd, Information measures, effective complexity, and total information, Complexity 2 (1996) 44–52

[13] J. Jost, External and internal complexity of complex adaptive systems, Theory in Biosciences 123 (2004) 69–88

[14] T. Kahle, Neighborliness of Marginal Polytopes, Contributions to Algebra and Geometry 51 (1) (2010) 45–56

[15] A. S. Klyubin, D. Polani, and C. L. Nehaniv, Tracking information flow through the environment: Simple cases of stigmerg, in J. Pollack, editor, Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems, pages 563–568. MIT Press, 2004

[16] F. Matúš, N. Ay, On maximization of the information divergence from an exponential family, Proceedings of WUPES'03 (ed. J. Vejnarová), University of Economics Prague (2003) 199–204

[17] G. Montúfar, N. Ay, Refinements of Universal Approximation Results for Deep Belief Networks and Restricted Boltzmann Machines, Neural Computation 23 (5) (2011) 1306–1319

[18] G. Montúfar, K. Zahedi, N. Ay, A Theory of Cheap Control in Embodied Systems, submitted, arXiv:1407.6836

[19] G. Montúfar, N. Ay, K. Zahedi, Expressive Power of Conditional Restricted Boltzmann Machines, submitted, arXiv:1402.3346

[20] G. Montúfar, J. Rauh, N. Ay, Maximal Information Divergence from Statistical Models Defined by Neural Networks, Geometric Science of Information 2013, arXiv:1303.0268

[21] G. Montúfar, J. Rauh, N. Ay, Information Geometry of Polytopes, Entropy 16 (6) (2014) 3207–3233

[22] G. Montúfar, J. Rauh, N. Ay, Expressive Power and Approximation Errors of Restricted Boltzmann Machines, Proceedings of NIPS 2011

[23] B. A. Olshausen, D. J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images, Nature 381 (1996) 607–609

[24] R. Pfeifer and J. C. Bongard, How the Body Shapes the Way We Think: A New View of Intelligence, The MIT Press (Bradford Books), Cambridge, MA, 2006

[25] J. Rauh, N. Ay, Robustness, canalising functions, and systems design, Theory in Biosciences (2013), doi: 10.1007/s12064-013-0186-3

[26] J. Rissanen, Stochastic complexity in statistical inquiry, World Scientific, Singapore, 1989

[27] N. Le Roux and Y. Bengio, Representational power of restricted Boltzmann machines and deep belief networks, Neural Computation 20 (6) (2008) 1631–1649

[28] V. N. Vapnik, Statistical learning theory, J. Wiley, New York, 1998

[29] K. Zahedi, N. Ay, Quantifying Morphological Computation, Entropy 15 (5) (2013) 1887–1915, doi: 10.3390/e15051887

[30] K. Zahedi, N. Ay, R. Der, Higher coordination with less control – A result of information maximisation in the sensori-motor loop, Adaptive Behavior 18 (2010) 338–355