

# Outsourcing Control Requires Control Complexity

---

Carlotta Langer\*

Hamburg University of Technology  
Institute for Data Science  
Foundations  
carlotta.langer@tuhh.de

Nihat Ay

Hamburg University of Technology  
Institute for Data Science  
Foundations  
Santa Fe Institute  
Leipzig University  
Faculty of Mathematics and  
Computer Science

**Abstract** An embodied agent influences its environment and is influenced by it. We use the sensorimotor loop to model these interactions and quantify the information flows in the system by information-theoretic measures. This includes a measure for the interaction among the agent's body and its environment, often referred to as morphological computation. Additionally, we examine the controller complexity, which can be seen in the context of the integrated information theory of consciousness. Applying this framework to an experimental setting with simulated agents allows us to analyze the interaction between an agent and its environment, as well as the complexity of its controller. Previous research revealed that a morphology adapted well to a task can substantially reduce the required complexity of the controller. In this work, we observe that the agents first have to understand the relevant dynamics of the environment to interact well with their surroundings. Hence an increased controller complexity can facilitate a better interaction between an agent's body and its environment.

---

## Keywords

Integrated information, morphological computation, em-algorithm, information theory, information geometry

---

## I Introduction

Every embodied agent, whether it is an animal, a human, or a robot, exists in constant interaction with its environment. The morphology of an agent's body has a significant impact on the nature of this interaction. The authors of the book *How the Body Shapes the Way We Think: A New View of Intelligence* (Pfeifer & Bongard, 2006) emphasize the importance of this interaction and the influence of it on the structure of the control architecture, that is, the brain of an agent. Pfeifer and Gómez (2009) expressed this notion more precisely in the following way: "There is a kind of trade-off or balance: the better the exploitation of the dynamics, the simpler the control, the less neural processing will be required" (p. 80). This suggests that the way an agent interacts with its environment has an impact on the complexity of its control architecture. Our previous research with simulated agents confirmed this intuition (Langer & Ay, 2021). There we observed that a better interaction with the environment reduces the necessity for a complex controller. This relationship insinuates that a sufficiently well-designed morphology might make a complex control architecture superfluous.

In this work, we extend the framework to include the process of learning a new task. Thereby we are able to observe that a high controller complexity can facilitate a better interaction with the environment, captured by a measure called *morphological computation*. Furthermore, agents with a simplified control architecture seem to be almost unable to learn a good strategy. Hence we conclude that agents need an increased controller complexity to learn and that both concepts, the controller complexity and morphological computation, influence each other.

---

\* Corresponding author.

In the next section, we describe the historical background and outline, before we discuss the intuition in more detail.

### 1.1 Historical Background and Outline

In this work, we analyze the dynamics of the information flows in simple, simulated agents. Here we apply methods from information theory, which is based on the mathematical theory of communication introduced by Shannon (1948). There the author quantifies the properties of a communication channel, and we use related information-theoretic measures to quantify information flows among an agent's body, controller, and environment. These information flows are modeled by the sensorimotor loop, similar to the approaches applied by, for example, Polani and Möller (2009), Polani et al. (2007), and Tishby and Polani (2011).

The sensorimotor loop reflects the interactions among the elements of the sensors  $S$ , the actuators  $A$ , and the controller  $C$  and can be translated to probability distributions that define the agents' behavior. In our experiments, the agents are faced with the task of moving through a race-track environment without touching the walls. This is discussed in more detail in the Setting of the Experiment section. Additionally, some agents, the internal-world-model agents, are also equipped with an internal prediction of the next sensory state  $S'$ , and we call the mechanism that generates this state the *internal world model*. Such an internal world model was also used by Ay and Zahedi (2013). This approach allows us to analyze the information flows among the different parts of the agents, and especially the prediction, in detail. An overview of the different agents can be found in Figure 4.

The applied learning algorithm is defined in the Learning section, and it is based on a modification of the em-algorithm, a well-known information-geometric algorithm (Amari, 1995; Csiszár & Tszunády, 1984). We combine two different instances of this algorithm to alternate between optimizing the agent's behavior and updating its world model. The behavior is optimized by maximizing the likelihood of a goal variable, and this follows the reasoning of the approach described by Attias (2003) and further analyzed by Toussaint (2009) and Toussaint et al. (2006), called *planning as inference*. We also use this optimization in Langer and Ay (2021).

While the agents learn, we calculate various information-theoretic measures, defined in the Measures of the Information Flow section. One important aspect is to assess the complexity of the controller, which we determine using two different measures that quantify distinct mechanisms of the controller. There exist various approaches to complexity. In this work, we consider a system to be complex if it is more than the sum of its parts. Hence the first measure that contributes to the controller complexity quantifies how much information integration exists between two parts of the controller. If we are able to divide the controller into two distinct parts without loss of functionality, then we call it split and not complex.

This measure can be seen in the context of the integrated information theory (IIT) of consciousness, originally proposed by Tononi. The core idea of IIT is that the level of consciousness of a system can be equated with the amount of information integration among different parts of it. This theory developed rapidly from a measure for brain complexity (Tononi et al., 1994) toward a broad theory of consciousness (Barbosa et al., 2021; Oizumi et al., 2014; Tononi, 2004). Hence there exist various types of integrated information measures depending on the version of the theory on which these measures are based and the setting in which they are defined. Here we use the information-geometric measure that we propose in Langer and Ay (2020) as a measure for the controller complexity. Thereby we follow the suggestion by Mediano et al. (2022) to adopt a more pragmatic point of view on integrated information measures.

Additionally, we calculate a measure of synergy of the internal world model to assess the controller complexity. This internal world model predicts the next sensory state and is vital for finding an optimal behavior. Here we measure the importance of the interplay between the different information flows going to the internal world model, and we call this the *synergistic prediction*.

The term *synergistic* suggests a relation between this measure and the context of the partial information decomposition of random variables. There the goal is to decompose the information that

a set of variables holds about a target variable into separate, non-negative terms, namely, into redundant, synergistic, and unique information, introduced by Williams and Beer (2010). There exist different definitions of these terms, for instance, the BROJA partial information decomposition in the case of two input variables, defined by Bertschinger et al. (2014). Using a similar approach to synergy as we apply here leads to a definition of unique information in Ghazi-Zahedi (2019). Alternatively, a measure for representational complexity of feed-forward networks is discussed by Ehrlich et al. (2024). This quantifies how much of a system needs to be observed simultaneously to access a particular piece of information.

In Langer and Ay (2021), we compare the controller complexity of an agent, in this case, given only by the integrated information, with its morphological computation. Here the concept of *morphological computation* describes the reduction of computational cost for the controller that results from the interaction of the agent's body with its environment. One example where morphological computation is applied is the field of soft robotics. There the softness of the robots' bodies leads to a lower computational cost when they, for example, grab fragile objects (Ghazi-Zahedi et al., 2017; Nakajima et al., 2013, 2014). Different understandings of morphological computation are discussed by Ghazi-Zahedi (2019) and Müller and Hoffmann (2017). Auerbach and Bongard (2014) analyzed simulated evolving agents and concluded that the complexity of the morphology of an agent depends on its environment. In the field of embodied artificial intelligence, the cheap design principle, formulated by Pfeifer and Bongard (2006), states that a robot's body should be constructed in a way that best exploits the properties of the environment. This should lead to a simpler control architecture. The cheap design principle is discussed in the context of universal approximations by Montúfar et al. (2015).

We confirm this intuition in Langer and Ay (2021) in experiments with simulated agents, where the comparison between the controller complexity and morphological computation leads to the result that they are inversely correlated. On one hand, this is intuitive, since the more the agent relies on the interaction of its body with the environment to solve a task, the less involvement of the controller is needed. On the other hand, this leads to the problem that now embodied intelligence is correlated with reduced involvement of the brain. If the morphology of an agent's body is intelligent enough, would it need a control architecture at all?

Here we want to present an additional perspective by considering the challenge of learning to perform a task. This entails updating an internal world model to predict the outcome of one's actions. Hence we measure the controller complexity not only via the integrated information but also by the complexity of the internal world model. We hypothesize that a learning process requires the agent to highly integrate the available information, hence that learning requires an increased controller complexity. Edlund et al. (2011) concluded that integrated information increases with the fitness of evolving agents. Albantakis et al. (2014) increased the complexity of the environment, which led to higher integrated information, and Albantakis and Tononi (2015) observed that high integrated information, benefits rich dynamical behavior. All these results are clear indications that a high information integration in the controller is beneficial for an embodied agent that is faced with a task.

Note that the complexity measures in the context of integrated information focus on the mechanistic structure of the information flow inside the controller and the internal world model, not the actual quality of the internal world model. An alternative perspective on assessing an internal model is to measure how much of the environmental state the internal model captures—how much of the environment it represents—giving rise to internal representations. Ashby (1956) postulated that the number of internal states in a controller needs to be greater than or equal to the system being controlled in order to be stable, hence this defines a lower bound on the size of a representation. The importance of predicting the next sensory states via an internal model is discussed, for instance, by Clark (2015).

The necessity of representing the environment for an artificial agent was called into question by Brooks (1991). This point of view and further criticism toward the representationalist approach are discussed in detail by Clark and Toribio (1994). A thorough introduction to the history of

representations can be found in Marstaller et al. (2013), where the authors define the representation explicitly as the information about the environment encoded in the internal states that goes beyond the information in the sensors. They show that this measure increases with the fitness of simulated agents equipped with an evolutionary algorithm. Aside from the different type of measure, our approach focuses on understanding how individual agents learn a new task in their lifespan, not on a population level.

Using the simulated learning agents, we first consider the results of a type that does not need to form an internal world model: the ideal-world-model agents. These agents have access to a sampled, external world model that describes their experiences instantaneously and accurately. In this utopian situation, agents do not require a complex controller to learn, and they behave mainly through reactive mechanisms, as long as their world model is accurate. In contrast, the internal-world-model agents, the ones that have to learn their internal world models, require an increased controller complexity to successfully learn. Once their world model is accurate, the integrated information value decreases, because the agent can then make use of the interactions with its environment, measured by morphological computation.

We summarize the intuition and main results of our experiments in the next section.

## 1.2 Intuition and Main Results

Learning a new task and adapting to changes in the world poses a difficult challenge. An important aspect of this is to predict the outcomes of one's actions. We theorize that even for seemingly easy situations, in which agents can manage without much involvement of the brain, learning the best behavior requires complex computations in the controller. We illustrate this in the following example.

Consider a child learning to ride a bike. Nearly every task the child has learned previously, for example, walking, speaking, or drawing, becomes harder when the child tries to do it fast. So the child expects that moving slowly will lead to the best outcome. According to its understanding of the world, its world model, riding a bike slowly is easier than doing it fast. Unfortunately, speed is required to stabilize a bike. The child is working with an inaccurate world model. So before the child can learn to ride a bike, it has to assess the information from its experiences and understand that faster can mean easier. It has to update its world model to learn and to be able to use the world in an optimal way.

To analyze these dynamics, we closely examine the information flows in learning agents. In Figure 1, we depict a sketch of an agent interacting with its environment, a fork, and highlight the different information flows that we analyze in this work. The agent perceives its environment through its eye, and we quantify the importance of the information flow from the sensors to the controller by a measure called *sensory information*. We assess the complexity of the controller by two different measures, namely, *integrated information* and *synergistic prediction*, both of which are described in more detail later. The information flow from the controller to the actuators, which then determine the actions of the agent, is measured by *control*. Last, the interaction among an agent's body and

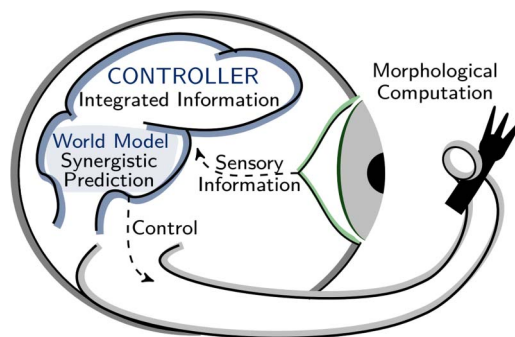


Figure 1. An agent interacting with its environment and the different measured information flows.

the environment, which reduces the computational cost for the controller, is called *morphological computation*. In the sketch of Figure 1, this is given by an octopus-like arm holding a fork.

We are especially interested in the controller complexity, quantified by two measures that assess distinct mechanisms of the controller. The first measure can be seen in the context of the integrated information theory of consciousness, introduced in the previous section, and it quantifies the information integration among parts of the controller. Additionally, the second measure assesses the complexity of the agent's internal world model, which predicts the next sensory state. It is called *synergistic prediction*. Both measures follow the notion that a system is complex, if it is more than the sum of its parts.

We use simple simulated agents and observe how the complexity of the controller develops during the learning process. Our first conclusion can be summarized as follows:

Conclusion 1: An agent that understands its environment, meaning that it has an accurate world model, exhibits a higher morphological computation and a lower controller complexity compared to agents with an inaccurate world model. The better an agent understands its environment, the more it can exploit the interactions between body and environment, and the less controller complexity is needed.

This conclusion is supported by the following observations. At first, we analyze agents that do not have to learn an internal world model. Instead, each of these agents is able to access an external world model, which samples the dynamics of the immediate environment of the agent. This then accurately describes the agents' experiences; it is ideal, hence we refer to these agents as *ideal-world-model agents*. We observe that they need next to no involvement of the controller and that the interaction with the environment, referred to as morphological computation, increases with the accuracy of the world model. At the same time, the influence of the controller on the agents' behavior is high for an inaccurate world model and decreases as the quality of the world model improves.

Furthermore, we refer to agents that have to learn an internal world model as *internal-world-model agents*. They initially have a high controller complexity, and then this value decreases, if they are successful. Hence we conclude earlier that the agents first have to learn the correct world model before they are able to optimally utilize the interaction of their bodies with the environment, which in turn leads to a lower controller complexity. Moreover, this theory is supported by the result that unsuccessful agents have a constantly high controller complexity and a lower morphological computation compared to the successful agents.

Additionally, we analyze agents with a simplified control architecture for which the ability to integrate information is inhibited. Hence the controller of these agents is divided into two unconnected parts; they have an integrated information of zero, and we call them *split internal-world-model agents*. The controller complexity of these agents is determined solely by the second measure, assessing the internal world model, and they perform noticeably worse compared to complete internal-world-model agents. The few successful split internal-world-model agents have a complex internal world model, which leads us to the following conclusion:

Conclusion 2: To successfully learn, the agents have to combine information from different sources. This leads to an increased controller complexity either in the form of integrated information or in the prediction process given by the internal world model.

In the next section, we introduce the experiments and the agents in more detail.

## 2 Materials and Methods

### 2.1 Setting of the Experiment

In our experiment, we analyze the information flows of simplistic, two-dimensional, acting agents. An agent consists of a round body with a radius of 0.55 unit length, a small tail, and two binary

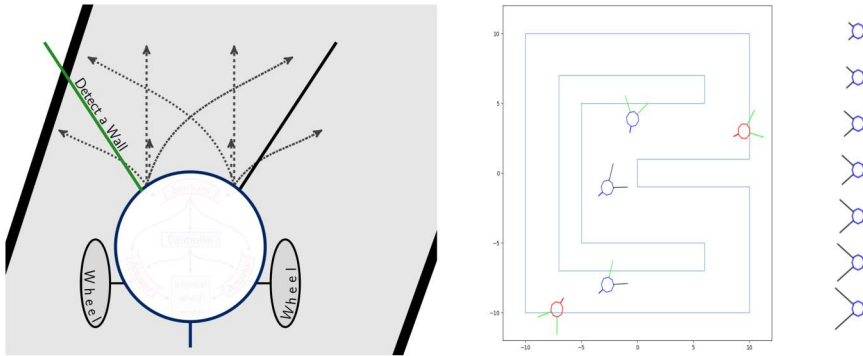


Figure 2. (left) A two-wheeled agent and its four movements in its environment. (middle) Five different agents in their environment. (right) Possible sensor lengths from 0.5 at the top to 2 at the bottom.

sensors. The tail simply marks the back of the agent and has no influence on its behavior. Two range sensors are visualized in Figure 2 (left) as lines that are green when they detect a wall, and black otherwise. We vary the reach of these sensors, as discussed in more detail subsequently. The agents can be thought of as two-wheeled robots, as sketched in Figure 2 (left). Each wheel can spin either fast or slow, which leads to four different movements, which are fast forward ( $\approx 0.6$  unit length per step), slow forward ( $\approx 0.2$  unit length per step), and left and right (with  $\approx 14^\circ$  and a speed of 0.4 unit length per step).

Five of these agents are depicted in Figure 2 (middle) on the racetrack in which they have to move. Whenever the body of an agent touches a wall, the agent gets stuck. This means that it can only turn on the spot and will not be able to move away unless both sensors no longer detect a wall. The implementation and a video of this movement can be found in Langer (2022).

Additionally, we vary the length of the sensors from 0.5, depicted in Figure 2 (at the top of the right panel), to a sensor length of 2, shown in Figure 2 (at the bottom of the right panel), with increments of 0.25. Varying the length of the sensors directly influences the amount of information an agent receives about the world, and hence it can influence the quality of the interaction of the agent with its environment. Therefore this has an impact on the potential morphological computation. Müller and Hoffmann (2017) call this *morphology facilitating perception* and discuss its relationship to other types of morphological computation in more detail.

## 2.2 The Agents and the World Models

An agent is modeled by a discrete multivariate, time-homogeneous Markov process, denoted by  $(X_t)_{t \in \mathbb{N}} = (S_t, A_t, C_t)$  with the state space  $\mathcal{X} = \mathcal{S} \times \mathcal{A} \times \mathcal{C}$ . The variable  $S_t$  entails the two binary sensors that detect a wall and a binary variable encoding whether the agent is touching a wall. The node  $A_t$  includes two binary actuators, and  $C_t$  includes two binary controller nodes. Additionally, in the case of the internal-world-model agents, the variable  $S'_t$  describes the internal prediction of the next sensor state and hence consists of three binary variables. The connections among these variables are sketched in Figure 3 (left).

The elements of the internal-world-model agents are connected according to the graph in Figure 3 (right). An introduction to graphical models is given by Lauritzen (1996). We depict only one node for each  $S, S', A,$  and  $C$  in the figures to increase clarity. The factorization of the corresponding probability distribution is given by

$$\begin{aligned}
 P(X_t, X_{t+1}, S'_{t+1}) &= P(S_t, A_t, C_t)P(S'_{t+1}|A_t, C_t)P(S_{t+1}|S_t, A_t) \\
 &\quad \prod_{j=1}^2 P(C_{t+1}^j|C_t, S_{t+1}) \prod_{i=1}^2 P(A_{t+1}^i|S_{t+1}, C_{t+1}) \tag{1}
 \end{aligned}$$

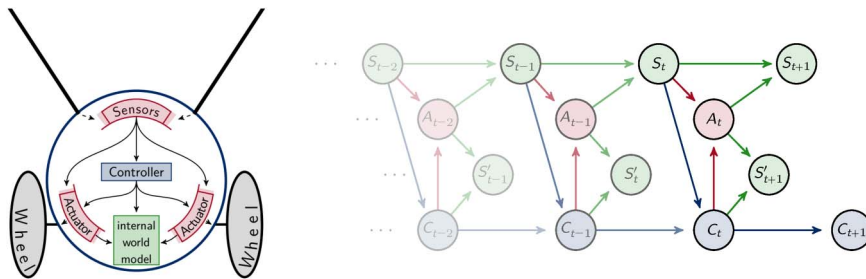


Figure 3. (left) Architecture of an internal-world-model agent. (right) Sensorimotor loop of this agent.

Because  $S'_t$  is a prediction of  $S_t$ , it is made of the same substrate, hence the state space of  $S'_t$  is  $\mathcal{S}$ . The difference between  $S_t$  and  $S'_t$  lies in the mechanism with which they are generated. The node  $S_t$  is influenced by the information from  $S_{t-1}$  and  $A_{t-1}$ . These are indirect influences, because in this case, the information flows through the environment. The role of the environment is discussed in more detail by Langer and Ay (2021).

The conditional distribution,  $P(S_{t+1}|S_t, A_t)$ , is called a *world model* by Montúfar et al. (2015) and Zahedi et al. (2010). The internal prediction  $S'_t$  is generated by  $P(S'_{t+1}|A_t, C_t)$  and is also named a *world model* by Ay and Zahedi (2013, 2014). To prevent confusion, we refer to  $P(S'_{t+1}|A_t, C_t)$  as an *internal world model* and to  $P(S_{t+1}|S_t, A_t)$  as an *ideal world model*. We use the latter term because this distribution is defined by sampling the individual past experiences of the agents, hence the ideal world model always represents the agents' experiences accurately.

In total, we analyze the behavior of four types of agents, summarized in Figure 4. The first distinction among the agents is between those with a complete controller, depicted in Figure 4 (left), and agents with a simplified controller, which are called split agents and are depicted in Figure 4 (right). The latter agents are not able to integrate information between the controller nodes, because the controller node  $C_{t+1}^i$  receives information only from  $C_t^i$  and not from  $C_t^j, i, j \in \{1, 2\}, i \neq j$ .

Second, we differentiate between agents with and without an internal world model. The agents in Figure 4 (top) have no internal world model. These agents have direct access to their sampled, ideal world model, and they are called ideal-world-model agents, whereas the internal-world-model agents, depicted in Figure 4 (bottom), have to learn their internal world models.

### 2.3 Learning

In our previous publication (Langer & Ay, 2021), we used the concept of *planning as inference* to optimize the agents' behavior, and we will apply the same algorithm here in the case of the ideal-world-model agents. In this method, the conditional distributions are optimized with respect to a goal variable by using the em-algorithm. This is a well-known information-geometric algorithm that is guaranteed to converge, but might converge to a local minimum (Amari, 1995; Amari & Nagaoka, 2007). This algorithm minimizes the difference between two sets of probability distributions by iteratively projecting onto them.

In the case of the internal-world-model agents, we have two goals. First, we want to optimize the distributions determining the behavior,  $P(C_{t+1}|S'_{t+1}, C_t)$  and  $P(A_{t+1}|S'_{t+1}, C_{t+1})$ , such that the probability of touching the wall after the next movement is as low as possible. At the same time, the internal world model,  $P(S'_{t+1}|C_t, A_t)$ , should be close to the actual, ideal world model,  $P(S_{t+1}|S_t, A_t)$ . This second goal is important, because otherwise, the optimization of the behavior would use faulty assumptions, leading to a failure of the agent. In the example in the Introduction section, this would be the child trying to learn to ride a bike while going as slowly as possible. Hence the two world models should result in similar predictions.

So we modify the em-algorithm to alternate between optimizing the agent with respect to the goal, on one hand, and with respect to the difference between world models, on the other hand. Details of this optimization are given in the appendix.

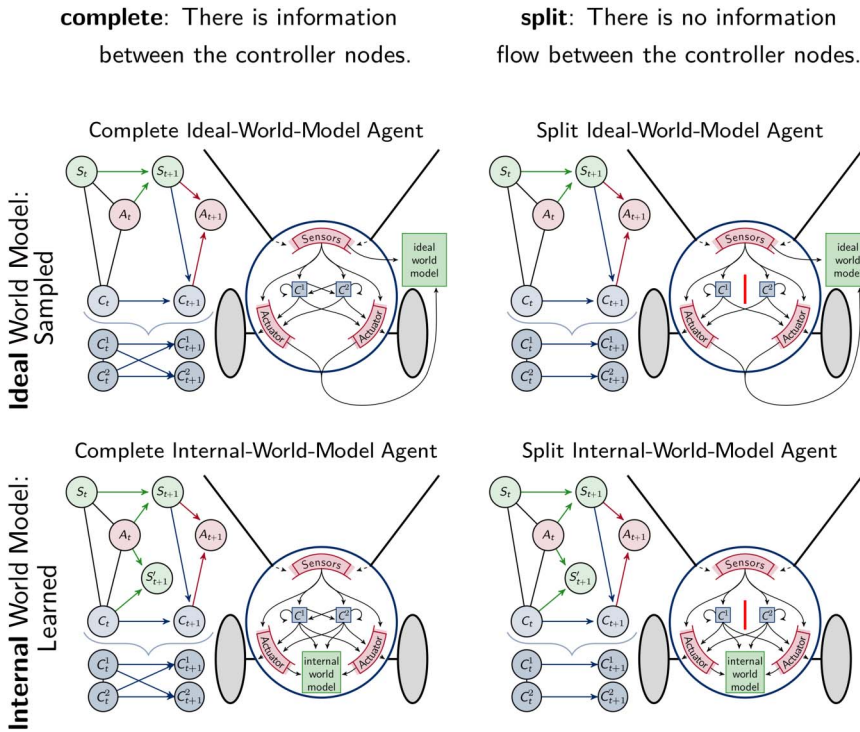


Figure 4. (top) Connections of the complete and split ideal-world-model agents. (bottom) Complete and split internal-world-model agents.

Note that the controller has only two binary variables, whereas  $\mathcal{S}$  consists of three. Therefore merely copying the information from the sensors is not a viable strategy for the agents. Even though we are studying simple agents here, this is a natural setting compared with human perception. We, as humans, do not consciously perceive every detail of our environments; instead, we learn to distinguish between important and irrelevant information.

### 2.4 Measures of the Information Flow

We measure the importance of an information flow by calculating the difference between the actual distribution and the closest distribution without the information flow in question. In Figure 5, we emphasize the measured connection by a dashed arrow. The set of distributions without this information flow is called a *split system*. More precisely, the measure in case of a split system  $M$  is defined in the following way.

**Definition** Let  $M$  be a set of positive probability distributions on a state space  $\mathcal{Z}$ , referred to as a split system. Then we define the measure  $\Psi_M$  by minimizing the Kullback–Leibler divergence between the split system  $M$  and the full distribution  $P$ :

$$\Psi_M = \inf_{Q \in M} D_{\mathcal{Z}}(P \parallel Q) = \inf_{Q \in M} \sum_{z \in \mathcal{Z}} P(z) \log \frac{P(z)}{Q(z)} \tag{2}$$

Most of the discussed measures have a closed-form solution and can be written in the form of sums of conditional mutual information terms. The conditional mutual information,  $I(Z_1; Z_2 | Z_3)$ , is defined by

$$I(Z_1; Z_2 | Z_3) = \sum_{z_1, z_2, z_3} P(z_1, z_2, z_3) \log \frac{P(z_1, z_2 | z_3)}{P(z_1 | z_3)P(z_2 | z_3)} \tag{3}$$

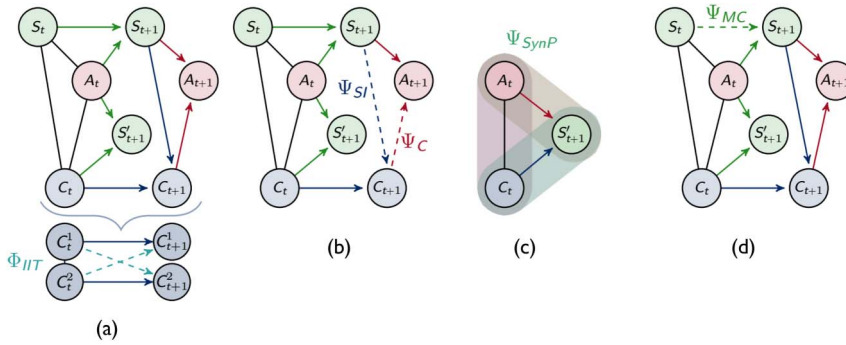


Figure 5. Graphs corresponding to the split systems in case of (a)  $\Phi_{IIT}$ , (b)  $\Psi_{SI}$  and  $\Psi_C$ , (c)  $\Psi_{SynP}$ , and (d)  $\Psi_{MC}$ .

and can be interpreted as follows: If  $I(Z_1; Z_2|Z_3) = 0$ , then  $Z_1$  is independent of  $Z_2$  given  $Z_3$ ; hence this quantifies the connection between  $Z_1$  and  $Z_2$ , given the influence of  $Z_3$ .

### 2.4.1 Controller Complexity

We assess the controller complexity using two different concepts that refer to different parts of the controller. First we discuss the measure corresponding to the integrated information, before we quantify the complexity of the internal world model.

**Integrated information.** There exist various types of integrated information measures, as discussed in the Introduction section. The approach we are using here was defined in our previous publication (Langer & Ay, 2020; also applied in Langer & Ay, 2021). There we quantify how much information gets integrated among different controller nodes across different points in time, as depicted in Figure 5(a). The minimization described in Equation 2 results in the following solution, as shown by Langer and Ay (2020):

$$\Phi_{IIT} = \sum_j I(C_{t+1}^j; C_t^{\setminus j} | C_t^j, S_{t+1}) \tag{4}$$

In our case, we have only two binary controller nodes, hence  $J = \{1, 2\}$ . Note that the split agents do not have these connections, leading to  $\Phi_{IIT} = 0$ .

The importance of the integrated information for the behavior of an agent also depends on the information flowing to and from the controller, as observed by Langer and Ay (2021). This is quantified by the two following measures, namely, sensory information and control.

We assess the importance of the information flow from the sensory nodes to the controller nodes by a measure called *sensory information*. The graphical representation of the split system is depicted in Figure 5(b) and the closed-form solution of this measure is

$$\Psi_{SI} = \sum_j I(C_{t+1}^j; S_{t+1} | C_t) \tag{5}$$

If this value is zero, then the controller nodes do not depend on the sensory input and therefore cannot make any behaviorally beneficial contributions.

Additionally, the strength of the connection from the controller nodes to the actuator nodes is assessed by a measure that we call *control*,  $\Psi_C$ :

$$\Psi_C = \sum_i I(A_{t+1}^i; C_{t+1} | S_{t+1}) \tag{6}$$

An agent with a controller that has no influence on the actuator at all has  $\Psi_C = 0$ .

The combination of these three measures quantifies the impact of the integrated information on the behavior of the agent. This is called *effective information integration* and is defined as the product

$$\Phi_{\text{EII}} = \Phi_{\text{IIT}} \cdot \Psi_{\text{SI}} \cdot \Psi_C \quad (7)$$

**Internal world model.** We analyze the internal world model  $P(S'_{t+1}|C_t, A_t)$  by calculating the importance of the interplay between the influences of  $A_t$  and  $C_t$  on  $S'_{t+1}$ .

This measure has no closed-form solution. Here we define a split system  $\mathcal{Q}$  as consisting of only the two-way interactions among the three variables, namely,  $\mathcal{Q}(A_t, C_t)$ ,  $\mathcal{Q}(A_t, S'_{t+1})$ , and  $\mathcal{Q}(C_t, S'_{t+1})$ , but without combined influence from  $(A_t, C_t)$  on  $S'_{t+1}$ . Hence we call this measure synergistic prediction,  $\Psi_{\text{SynP}}$ . The two-way interactions are highlighted in Figure 5(c). This is conceptually similar to the synergistic measure for morphological computation proposed by Ghazi-Zahedi, Langer, and Ay (2017), and we also use the iterative scaling algorithm to calculate this measure, as described there in section 2.5.

### 2.4.2 Morphological Computation

The concept of morphological computation describes the reduction of the necessary computation in the controller that results from the interaction of the agent's body with its environment. There exist various types of morphological computation and different measures for it (Müller & Hoffmann, 2017). We use the following formulation:

$$\Psi_{\text{MC}} = I(S_{t+1}; S_t | A_t) \quad (8)$$

This measures the information flow going from one sensory state to the next through the world, given the actuator state, as depicted in Figure 5(d). In Zahedi and Ay (2013), this was introduced as a measure for morphological computation, and in Ghazi-Zahedi (2019), in a comparison with other measures,  $\Psi_{\text{MC}}$  shows desirable properties.

## 3 Results

In this section, we discuss the results of our simulations. We used 1,000 random input distributions for each sensor length and each type of agent. All agents train for 20,000 steps, and the measures are calculated for 90 different points during these steps. More precisely, we apply the measures for the nine time points, namely, 50, 100, 200, 500, 1,000, 2,000, 5,000, 10,000, and 20,000, and additionally for nine equidistant time points between each of them, as well as nine equidistant time points between 0 and 50.

Additionally, we calculate the success rate (SR) of an agent by sampling for how many time points the agent is stuck at a wall during the 20,000 training steps. Hence a SR of 0.1 signifies that an agent was not stuck 10% of the steps. We then divide the agents into successful and unsuccessful ones based on their SRs. The best third of the complete internal-world-model agents perform above 16.8% and are called successful, while we refer to agents with a SR below 16.8% as unsuccessful. Dividing the agents in this way allows us only to call agents successful when the agents' SRs increased noticeably during learning.

In Figure 6, we can see the results of the measures for the controller complexity, namely, integrated information and synergistic prediction, as well as the morphological computation averaged over all successful internal-world-model agents after 20,000 steps.

We observe that the controller complexity and the morphological computation are inversely correlated. Therefore the results confirm our previous observation (Langer & Ay, 2021) that morphological computation and integrated information have an inverse relationship. Note that when

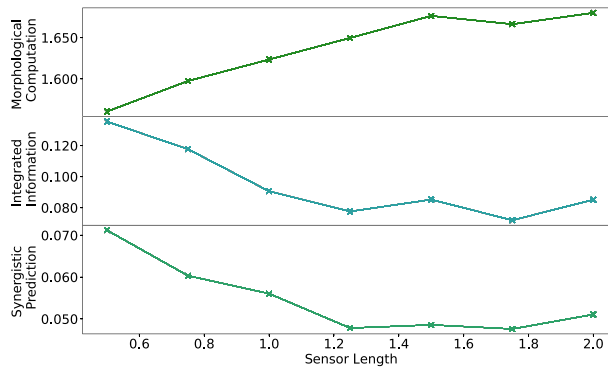


Figure 6. Morphological computation, integrated information, and synergistic prediction for the successful, complete internal-world-model agents.

the sensors are too long so that the agents almost always detect a wall, this additional information is no longer beneficial for the agents, and the morphological computation no longer increases, while the integrated information and synergistic prediction increase again.

This relationship leads to the question of why agents with a well-adapted morphology would need a complex control architecture. Wouldn't it be possible to build agents that are so well adapted to their environment that a simple controller suffices? There might be several reasons why a complex controller is necessary in general, despite this inverse correlation, as we discuss further in the Conclusion section.

Here we argue that an involvement of the controller is necessary because agents first have to learn how to interact with their environment, meaning that they have to build their world models.

### 3.1 The Ideal-World-Model Agents

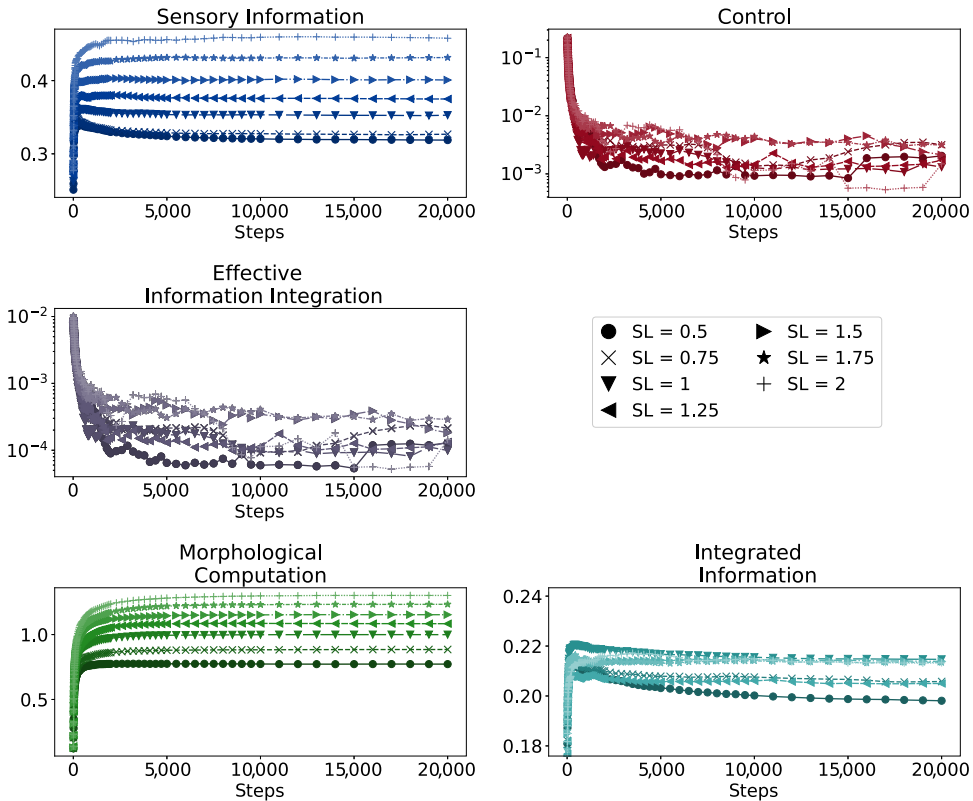
Now we discuss the results for the ideal-world-model agents that do not have to learn their world models because they have direct access to the sampled ideal world models. The best  $\approx 33\%$  of the ideal-world-model agents are the ones with a success rate higher than 61.5%, which we term the successful ideal-world-model agents. Hence the ideal-world-model agents perform overall much better than the internal-world-model agents, for which the best third only performs better than 16.8%. We depict the integrated information, sensory information, control, effective information integration, and morphological computation for the successful ideal-world-model agents in the first three rows of Figure 7.

The controller complexity, given here solely by the integrated information value due to the lack of an internal world model, seems not to change after the first few initial steps. In Langer and Ay (2021), we discussed that the importance of the controller complexity additionally depends on the sensory information and the control. While the sensory information increases with the sensor length, we can see the reason for the behavior of the integrated information in the results of  $\Psi_C$ , the measure for control. After the first steps, this measure is very close to 0, with an average value of 0.0021 at the 20,000th step. If  $\Psi_C = 0$ , then the controller has no influence on the behavior of the agent at all. It is easy to check that in this case, the information flow in the controller is no longer changed by the em-algorithm, because the controller has no influence on whether the agent is successful. This only holds for the ideal-world-model agents, because we apply the original em-algorithm here, not the modified one.

The effective information integration, depicted in the second row and first column of Figure 7, summarizes the behavior of the other three measures. This has a value close to zero, which shows that the controller complexity is nearly irrelevant for the behavior of the agent in this case.

Hence, for the ideal-world-model agents, a complex controller is not needed to learn to perform a task. In fact, split ideal-world-model agents, without the ability to integrate information, perform

Successful Agents



Unsuccessful Agents

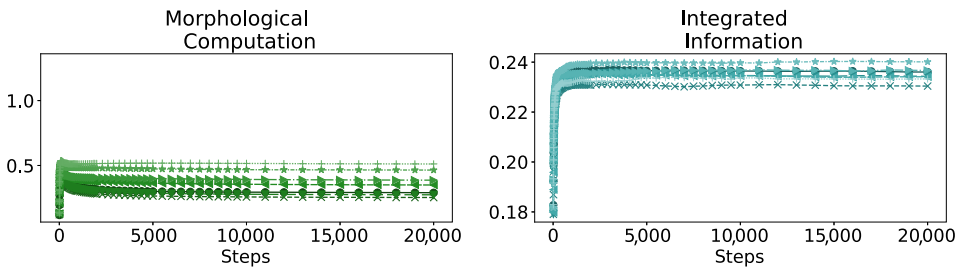


Figure 7. Sensory information, control, effective information integration, morphological computation, and integrated information for the successful ideal-world-model agents and morphological computation and integrated information for the unsuccessful ideal-world-model agents. SL = sensor length.

only slightly worse than complete ones. More precisely, the split ideal-world-model agents have an average success rate of 33.69%, compared to 33.83% in the complete case.

In this scenario, success depends, not on the complexity of the controller, but on the interaction of the agent with its environment. We therefore now directly compare the morphological computation and controller complexity of successful and unsuccessful ideal-world-model agents, depicted in the two bottom rows of Figure 7. The successful agents have a much higher morphological computation over all. The morphological computation measures how much the next sensor states depend on the last sensor states, given the actuator nodes, and is calculated using the ideal world

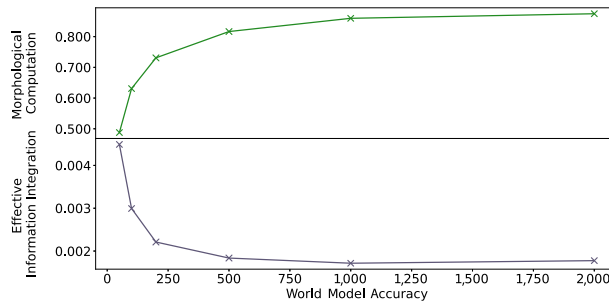


Figure 8. Morphological computation and effective information integration for the successful ideal-world-model agents with varying accuracy of the world model.

models. This means that the successful agents have found strategies to move in their environment and use the interaction with the environment in a way such that the next point in time is more predictable—more closely depends on the last sensory state—compared to the unsuccessful agents.

Additionally, the integrated information is overall higher in the case of the unsuccessful agents. There the agents have a lower morphological computation, and we again observe an inverse correlation between these two quantities. Previously we noted that the integrated information is not influenced by the em-algorithm after the first steps; however, the observation made here refers to the value that the algorithm reaches exactly during these first steps.

To conclude, if we have an ideal-world-model agent with access to its correct world model and with a morphology that is well adapted to its environment, then the ideal-world-model agent has no need for a complex control architecture—a brain.

To further examine this connection between the quality of the world model and the need for a complex controller, we additionally analyze agents that are only able to sample their ideal world models for a part of the total 20,000 steps. These agents sample the ideal world model and learn their behavior only up to a certain point. After that point, the world model stays fixed, and the agents have to use this, possibly inaccurate, world model to find the best behavior for the remainder of the 20,000 steps. We distinguish between nine different cases, namely, agents that sample the world model for 50, 100, 200, 500, 1,000, 2,000, 5,000, 10,000, or the full 20,000 steps.

In Figure 8, we highlight the relationship between morphological computation and effective information integration, defined in Equation 7, with respect to the accuracy of the world model on the  $x$  axis. There we display the arithmetic mean over the different sensor lengths after 20,000 steps. While the morphological computation increases with the accuracy of the world model, the effective information integration decreases.

In the Introduction section, we motivate the intuition behind these concepts using the example of a child learning to ride a bike. The better the child understands the dynamics of its environment, the more it can make use of them, and the faster it drives to stabilize the bike. Hence a better world model leads to a higher morphological computation, which then reduces the necessity for a complex controller.

This concludes the analysis of the accuracy of the ideal world model in relationship to the information flow inside the agents. In the next section, we discuss the internal-world-model agents that additionally have to learn their internal world models.

### 3.2 The Internal-World-Model Agents

Here we discuss the results for the internal-world-model agents, which have to learn the dynamics of the world that are relevant to the agents. We first focus on the measures for the controller complexity, called integrated information and synergistic prediction. Figure 9 depicts the success rate and all the information-theoretic measures for the successful and unsuccessful internal-world-model agents. Each line in these figures corresponds to a sensor length (SL) and is depicted with respect

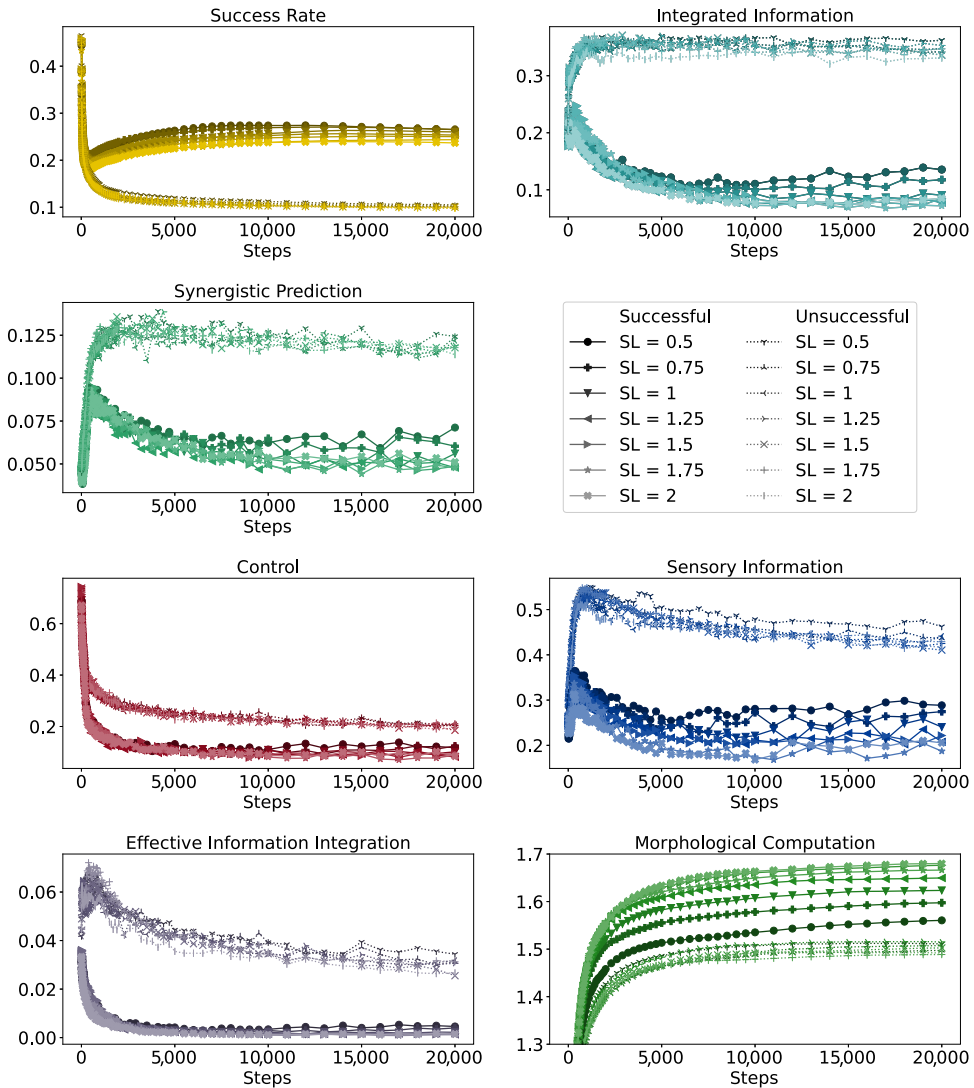


Figure 9. Success rate, integrated information, synergistic prediction, control, sensory information, effective information integration, and morphological computation for the successful and unsuccessful internal-world-model agents.

to the number of steps. The unsuccessful agents have an integrated information value of  $\approx 0.3-0.38$  and a synergistic prediction of  $\approx 0.1-0.14$ .

Now we compare these results to the integrated information value of the successful agents, and we can observe that there is first an increase in the integrated information and synergistic prediction values in the first 400 steps, then a strong decrease. After 20,000 steps, the integrated information value is roughly between 0.05 and 0.15, and the synergistic prediction lies between 0.04 and 0.08. Hence, in the case of the successful agents, the complexity of the controller reduces to a much lower value over time compared to the unsuccessful agents.

Following the observations of the previous section leads to the conclusion that a high controller complexity might be important as long as the agents have not been able to learn the correct world model. Without a correct world model, the agents are not able to find a strategy that would allow them to use their interaction with the environment optimally.

To further interpret these results in relation to the agents' learning behavior, we now discuss the values for the sensory information, control, and morphological computation. These first two give insights into the effect the integrated information has on the action of the agent and combined lead to the effective information integration. In the two bottom rows of Figure 9, we depict these four measures. The sensory information and control decrease with the number of steps taken for the successful, as well as for the unsuccessful, agents. However, there is a clear difference in the overall values of these measures, which leads to the effective information integration of the successful agents being  $\approx 0.002$ , whereas this value reaches on average 0.03 in the case of the unsuccessful agents. Hence the integrated information not only is higher for the unsuccessful agents, but also has more impact on the agents' behavior.

In the case of the morphological computation, we observe that the successful agents reach a higher morphological computation value, on average 1.64, compared to a value of 1.5 in the case of the unsuccessful agents.

These results support our hypothesis. A high controller complexity value seems to be important as long as the agents have not been able to learn to interact with their environment. Hence the morphological computation is lower for the unsuccessful agents, whereas the complexity and involvement of the controller are higher. Now the question remains whether a high controller complexity is really necessary for learning or just a by-product of the morphological computation being low. To clarify that point, we now look at the split internal-world-model agents, which have a simplified control architecture.

### 3.3 Comparing the Split and Complete Internal-World-Model Agents

The architecture of the split internal-world-model agents is depicted in Figure 4 (bottom right). These agents are not able to integrate information between their controller nodes, hence the complexity of the controller depends solely on the structure of the internal world model.

We divide these agents into successful and unsuccessful ones by applying the criterion for success from the internal-world-model agents of 16.8%. The split agents perform worse, and this does not lead to a one-third/two-third split. However, thereby we are able to directly compare complete internal-world-model and split agents with similar success rates.

First, we consider the average success rates of the split and complete internal-world-model agents. In addition, we compare them with the average success rate of agents that perform only random movements and do not learn at all. For random movement, the average success rate is  $\approx 7.95\%$ ; for complete internal-world-model agents, it is  $\approx 15.21\%$ , and for split internal-world-model agents, it is  $\approx 8.01\%$ .

The split internal-world-model agents perform on average barely better than the agents that move randomly. Note that there is also a considerable difference in the number of successful agents. Only  $\approx 2.1\%$  of split internal-world-model agents are successful compared to 33.3% of the complete ones.

In summary, the split agents perform only marginally better than agents that move purely at random, and only very few split agents are successful. This strongly supports the hypothesis that an increased controller complexity is necessary for learning.

Additionally, we now focus on the internal world model of the few successful agents. Here we compare the synergistic prediction of the successful, complete internal-world-model and the successful, split internal-world-model agents. The results are shown in Figure 10.

The synergistic prediction quantifies how important the interaction of both influences, from the actuators as well as from the controller nodes, is for the prediction. It is noticeable that the synergistic prediction is much higher for the successful agents that are not able to integrate information. This leads to the conclusion that for these split agents, the internal world model, and therefore the prediction process, has to combine the information from different sources and becomes much more complex. The complete internal-world-model agents are able to integrate the information directly between their controller nodes and do not need such a complicated world model to have a complex controller.

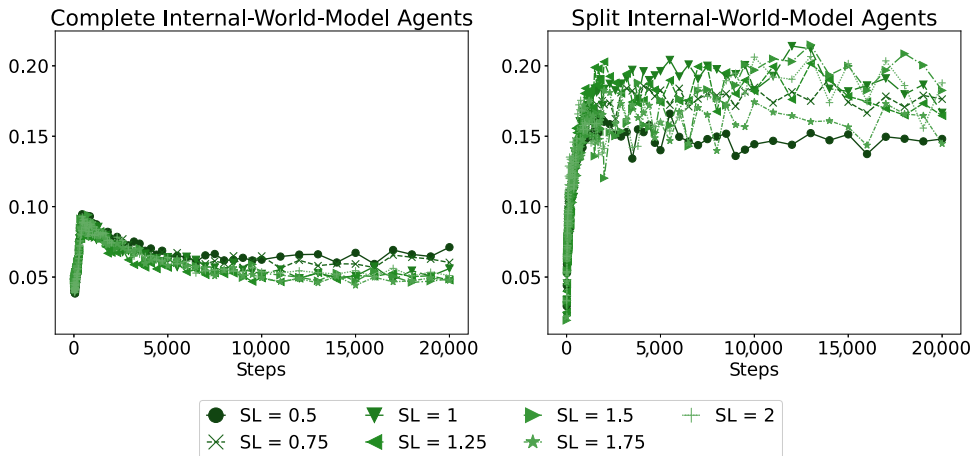


Figure 10. Comparison of synergistic prediction in case of (left) the successful internal-world-model agents and (right) the successful split internal-world-model agents.

## 4 Conclusion

In this article, we discuss the dynamics of morphological computation and controller complexity in learning, embodied artificial agents. These agents move inside a racetrack and learn not to touch the walls. Using this simplistic example, we are able to analyze the different information flows inside the agents and, especially, examine the process of predicting the next sensory state. As a training algorithm, we use an adapted em-algorithm that alternates between optimizing the behavior to reach a goal and updating the internal world model. This algorithm fits naturally into our framework, because it is information geometric in nature. Additionally, its geometric interpretation highlights the interplay between the goals of optimizing the behavior and the world model. However, for future work, we intend to analyze the influence that more biologically plausible learning algorithms have on the information flows inside the system.

The results of our experiment regarding the controller complexity and morphological computation support our previous publication (Langer & Ay, 2021), because we observe the inverse correlation between them. These previous results lead to the insinuation that agents with a highly adapted morphology might have no use for a complex control architecture. There are many possible ways to address this notion. It might be that our tasks are simply too easy to solve so that an agent truly only needs morphological computation to be successful. Another possibility is given by Pfeifer and Gómez (2009):

The more the specific environmental conditions are exploited—and the passive dynamic walker is an extreme case—the more the agent’s success will be contingent upon them.

Thus, if we really want to achieve brain-like intelligence, the brain (or the controller) must have the ability to quickly switch to different kinds of exploitation schemes either neurally, or mechanically through morphological change. (p. 80)

Hence the agents might have no need for a controller because they are faced with only a single task, namely, avoiding the walls of their environment. Furthermore, the nature of the task might be too simplistic, and it might have to require a higher order of understanding of its surroundings so that the agents truly need to process the information from the environment. Therefore we will develop this approach further to explore these possibilities and apply it to more involved settings.

Despite the simplicity of our example, we were able to offer an additional solution to the posed problem. We theorize that learning to predict the environment results in a necessity for a complex

controller. Ideal-world-model agents, which do not have to learn to predict their environment, do not require a complex controller at all, not even to learn our task. However, when their ability to form an accurate world model is restricted, the involvement of the control architecture increases.

The internal-world-model agents, on the other hand, show a necessity for an increased controller complexity in general. The controller complexity of the successful agents is first high, while the agents learn their world model, then decreases. We argue that this decrease could result from a rise in morphological computation that is facilitated by the correct world model. This is supported by the results for morphological computation, which are higher in the case of the successful agents. Hence the two quantities, the controller complexity and the morphological computation, influence each other.

Comparing the complete internal-world-model agents with the split ones, which have a simplified controller and are not able to integrate information, leads to the observation that the latter agents are not able to predict the next sensory state as well. The split internal-world-model agents perform on average only marginally better than completely randomly moving agents, and there is only a very small percentage of successful split internal-world-model agents. Hence learning requires an increased controller complexity.

Furthermore, the few successful, split internal-world-model agents have a more complex prediction process. This process itself combines the information from the controller and the actuator nodes to form a prediction of the next sensory state. This again supports the claim that an agent needs to integrate its available information to learn. In this case, the complex process is not directly between the controller nodes but inside the internal world model.

## Funding Information

The authors acknowledge funding by Deutsche Forschungsgemeinschaft Priority Programme “The Active Self” (SPP 2134).

## References

- Albantakis, L., Hintze, A., Koch, C., Adami, C., & Tononi, G. (2014). Evolution of integrated causal structures in animats exposed to environments of increasing complexity. *PLoS Computational Biology*, *10*(12), Article e1003966. <https://doi.org/10.1371/journal.pcbi.1003966>, PubMed: 25521484
- Albantakis, L., & Tononi, G. (2015). The intrinsic cause-effect power of discrete dynamical systems—from elementary cellular automata to adapting animats. *Entropy*, *17*(8), 5472–5502. <https://doi.org/10.3390/e17085472>
- Amari, S. (1995). Information geometry of the EM and em algorithms for neural networks. *Neural Networks*, *8*(9), 1379–1408. [https://doi.org/10.1016/0893-6080\(95\)00003-8](https://doi.org/10.1016/0893-6080(95)00003-8)
- Amari, S., & Nagaoka, H. (2007). *Methods of information geometry*. American Mathematical Society. <https://doi.org/10.1090/mmono/191>
- Ashby, W. R. (1956). *An introduction to cybernetics*. Wiley. <https://doi.org/10.5962/bhl.title.5851>
- Attias, H. (2003). Planning by probabilistic inference. In C. M. Bishop & B. J. Frey (Eds.), *International workshop on artificial intelligence and statistics, 3–6 January 2003, Key West, Florida, USA* (pp. 9–16). PMLR.
- Auerbach, J. E., & Bongard, J. C. (2014). Environmental influence on the evolution of morphological complexity in machines. *PLoS Computational Biology*, *10*(1), Article e1003399. <https://doi.org/10.1371/journal.pcbi.1003399>, PubMed: 24391483
- Ay, N., & Zahedi, K. (2013). Causal effects for prediction and deliberative decision making of embodied systems. In Y. Yamaguchi (Ed.), *Advances in cognitive neurodynamics* (Vol. 3, pp. 499–506). Springer. [https://doi.org/10.1007/978-94-007-4792-0\\_67](https://doi.org/10.1007/978-94-007-4792-0_67)
- Ay, N., & Zahedi, K. (2014). On the causal structure of the sensorimotor loop. In M. Prokopenko (Ed.), *Guided self-organization: Inception* (pp. 261–294). Springer. [https://doi.org/10.1007/978-3-642-53734-9\\_9](https://doi.org/10.1007/978-3-642-53734-9_9)
- Barbosa, L. S., Marshall, W., Albantakis, L., & Tononi, G. (2021). Mechanism integrated information. *Entropy*, *23*(3), Article 362. <https://doi.org/10.3390/e23030362>, PubMed: 33803765

- Bertschinger, N., Rauh, J., Olbrich, E., Jost, J., & Ay, N. (2014). Quantifying unique information. *Entropy*, *16*(4), 2161–2183. <https://doi.org/10.3390/e16042161>
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, *47*(1–3), 139–159. [https://doi.org/10.1016/0004-3702\(91\)90053-M](https://doi.org/10.1016/0004-3702(91)90053-M)
- Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190217013.001.0001>
- Clark, A., & Toribio, J. (1994). Doing without representing? *Synthese*, *101*, 401–431. <https://doi.org/10.1007/BF01063896>
- Csiszár, I., & Tsunády, G. (1984). Information geometry and alternating minimization procedures. *Statistics and Decisions*, *1*, 205–237.
- Edlund, J. A., Chaumont, N., Hintze, A., Koch, C., Tononi, G., & Adami, C. (2011). Integrated information increases with fitness in the evolution of animats. *PLoS Computational Biology*, *7*(10), Article e1002236. <https://doi.org/10.1371/journal.pcbi.1002236>, PubMed: 22028639
- Ehrlich, D. A., Schneider, A. C., Priesemann, V., Wibral, M., & Makkeh, A. (2024). *A measure of the complexity of neural representations based on partial information decomposition*. Unpublished manuscript.
- Ghazi-Zahedi, K. (2019). Morphological intelligence. In *Concept Three: Quantifying Morphological Intelligence as Synergy of Body and Brain* (pp. 89–94). Springer. <https://doi.org/10.1007/978-3-030-20621-5>
- Ghazi-Zahedi, K., Deimel, R., Montúfar, G., Wall, V., & Brock, O. (2017). Morphological computation: The good, the bad, and the ugly. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 464–469). IEEE. <https://doi.org/10.1109/IROS.2017.8202194>
- Ghazi-Zahedi, K., Langer, C., & Ay, N. (2017). Morphological computation: Synergy of body and brain. *Entropy*, *19*(9), Article 456. <https://doi.org/10.3390/e19090456>
- Langer, C. (2022). *Learning to predict requires integrated information*. <https://github.com/CarlottaLanger/LearningRequiresIntInf>
- Langer, C., & Ay, N. (2020). Complexity as causal information integration. *Entropy*, *22*(10), 1107. <https://doi.org/10.3390/e22101107>, PubMed: 33286876
- Langer, C., & Ay, N. (2021). How morphological computation shapes integrated information in embodied agents. *Frontiers in Psychology*, *12*, 716433. <https://doi.org/10.3389/fpsyg.2021.716433>, PubMed: 34912262
- Lauritzen, S. L. (1996). *Graphical models*. Clarendon Press. <https://doi.org/10.1093/oso/9780198522195.001.0001>
- Marstaller, L., Hintze, A., & Adami, C. (2013). The evolution of representation in simple cognitive networks. *Neural Computation*, *25*(8), 2079–2107. [https://doi.org/10.1162/NECO\\_a\\_00475](https://doi.org/10.1162/NECO_a_00475), PubMed: 23663146
- Mediano, P. A. M., Rosas, F. E., Farah, J. C., Shanahan, M., Bor, D., & Barrett, A. B. (2022). Integrated information as a common signature of dynamical and information-processing complexity. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *32*(1), 013115. <https://doi.org/10.1063/5.0063384>, PubMed: 35105139
- Montúfar, G., Ghazi-Zahedi, K., & Ay, N. (2015). A theory of cheap control in embodied systems. *PLoS Computational Biology*, *11*(9), Article e1004427. <https://doi.org/10.1371/journal.pcbi.1004427>, PubMed: 26325254
- Müller, V. C., & Hoffmann, M. (2017). What is morphological computation? On how the body contributes to cognition and control. *Artificial Life*, *23*(1), 1–24. [https://doi.org/10.1162/ARTL\\_a\\_00219](https://doi.org/10.1162/ARTL_a_00219), PubMed: 28140632
- Nakajima, K., Hauser, H., Kang, R., Guglielmino, E., Caldwell, D. G., & Pfeifer, R. (2013). A soft body as a reservoir: Case studies in a dynamic model of octopus-inspired soft robotic arm. *Frontiers in Computational Neuroscience*, *7*, 91. <https://doi.org/10.3389/fncom.2013.00091>, PubMed: 23847526
- Nakajima, K., Li, T., Hauser, H., & Pfeifer, R. (2014). Exploiting short-term memory in soft body dynamics as a computational resource. *Journal of the Royal Society Interface*, *11*(100), 20140437. <https://doi.org/10.1098/rsif.2014.0437>, PubMed: 25185579
- Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS Computational Biology*, *10*(5), Article e1003588. <https://doi.org/10.1371/journal.pcbi.1003588>, PubMed: 24811198

- Pfeifer, R., & Bongard, J. (2006). *How the body shapes the way we think: A new view of intelligence*. MIT Press. <https://doi.org/10.7551/mitpress/3585.001.0001>
- Pfeifer, R., & Gómez, G. (2009). Morphological computation—Connecting brain, body, and environment. In B. Sendhoff, E. Körner, O. Sporns, H. Ritter, & K. Doya (Eds.), *Creating brain-like intelligence: From basic principles to complex intelligent systems* (pp. 66–83). Springer. [https://doi.org/10.1007/978-3-642-00616-6\\_5](https://doi.org/10.1007/978-3-642-00616-6_5)
- Polani, D., & Möller, M. (2009). Models of information processing in the sensorimotor loop. In F. Emmert-Streib & M. Dehmer (Eds.), *Information theory and statistical learning* (pp. 289–308). Springer. [https://doi.org/10.1007/978-0-387-84816-7\\_12](https://doi.org/10.1007/978-0-387-84816-7_12)
- Polani, D., Sporns, O., & Lungarella, M. (2007). How information and embodiment shape intelligent information processing. In M. Lungarella, F. Iida, J. Bongard, & R. Pfeifer (Eds.), *50 years of artificial intelligence* (pp. 99–111). Springer. [https://doi.org/10.1007/978-3-540-77296-5\\_10](https://doi.org/10.1007/978-3-540-77296-5_10)
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 623–656. <https://doi.org/10.1002/j.1538-7305.1948.tb00917.x>
- Tishby, N., & Polani, D. (2011). Information theory of decisions and actions. In V. Cutsuridis, A. Hussain, & J. G. Taylor (Eds.), *Perception-action cycle: Models, architectures, and hardware* (pp. 601–636). Springer. [https://doi.org/10.1007/978-1-4419-1452-1\\_19](https://doi.org/10.1007/978-1-4419-1452-1_19)
- Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5, Article 42. <https://doi.org/10.1186/1471-2202-5-42>, PubMed: 15522121
- Tononi, G., Sporns, O., & Edelman, G. M. (1994). A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences of the United States of America*, 91(11), 5033–5037. <https://doi.org/10.1073/pnas.91.11.5033>, PubMed: 8197179
- Toussaint, M. (2009). Probabilistic inference as a model of planned behavior. *Künstliche Intelligenz*, 23(3), 23–29.
- Toussaint, M., Harmeling, S., & Storkey, A. (2006). *Probabilistic inference for solving (PO) MDPs* (Technical Report No. 934). School of Informatics, University of Edinburgh.
- Williams, P. L., & Beer, R. D. (2010). *Nonnegative decomposition of multivariate information*. ArXiv. <https://doi.org/10.48550/arXiv.1004.2515>
- Zahedi, K., & Ay, N. (2013). Quantifying morphological computation. *Entropy*, 15(5), 1887–1915. <https://doi.org/10.3390/e15051887>
- Zahedi, K., Ay, N., & Der, R. (2010). Higher coordination with less control—a result of information maximization in the sensorimotor loop. *Adaptive Behavior*, 18(3–4), 338–355. <https://doi.org/10.1177/1059712310375314>

## Appendix: Learning the Strategy and the World Model

The learning algorithm applied to the internal-world-model agents in our experiments works by adapting the em-algorithm to incorporate two different goals. The em-algorithm is a well-known information-geometric algorithm that iteratively projects to two different sets of probability distributions and thereby reduces the KL-divergence between them (Amari, 1995; Amari & Nagaoka, 2007).

In this article, an agent learns inside the racetrack. Hence the realized states  $s_{t-1}$ ,  $a_{t-1}$ , and  $c_{t-1}$  are known at each step  $t$  and can be used. To that end, we need the following definitions.

Let  $P_{c_t}(C_{t+1}|S_{t+1})$  be the distribution of  $C_{t+1}$  conditioned on  $S_{t+1}$  and a fixed state  $c_t$ , meaning that  $P_{c_t}(c_{t+1}|s_{t+1}) := P(c_{t+1}|s_{t+1}, c_t)$  for all  $s_{t+1}, c_{t+1} \in \mathcal{S} \times \mathcal{C}$ .

The ideal-world-model agents are able to use the sampled world model for the prediction, whereas the internal-world-model agents make use of their internal world model to arrive at the internal prediction  $s'$ . Both types of agents can optimize the distributions  $P_{c_t}(c_{t+1}|s_{t+1})$  and  $P(a_{t+1}|s_{t+1}, c_{t+1})$  with  $s_{t+1}, a_{t+1}, c_{t+1} \in \mathcal{S} \times \mathcal{A} \times \mathcal{C}$ . In addition, the internal-world-model agents learn their internal world model given by the distribution  $P(S'_{t+2}|A_{t+1}, C_{t+1})$ .

Furthermore, we add Gaussian noise to the distribution  $P(A_{t+1}|S'_{t+1}, C_{t+1})$ , because the em-algorithm cannot gain a positive value again once it reaches a point where for some action  $a_{t+1}$  the equality  $P(a_{t+1}|s'_{t+1}, c_{t+1}) = 0$  holds.

Now we introduce the modified em-algorithm. First we define one set for optimizing with respect to the goal. Let  $S^3$  be the variable indicating whether the agent is touching a wall. Then  $s^3 = 1$  signifies that the agent is not touching a wall. Now let  $X'_{t+1} = (S'_{t+1}, A_{t+1}, C_{t+1})$ ; then the goal manifold consists of all those probability distributions for which it is certain that the agent will not touch a wall at point  $t + 2$ :

$$\mathcal{M}_G^P(x_t) = \{Q \in \mathcal{P}(\mathcal{X} \times \mathcal{S}) \mid Q(s^3_{t+2} = 1) = 1\} \quad (9)$$

The second set consists of all the distributions that factor according to the agents, meaning that they consist of all the possible agents, given the current world model:

$$\mathcal{M}_{\mathcal{A}}^P(x_t, \bar{P}) = \left\{ P \in \mathcal{P}^\circ(\mathcal{X} \times \mathcal{S}) \mid P(x_{t+1}, s_{t+2}) = \bar{P}_{a_t, c_t}(s_{t+1}) \prod_j P_{c_t}(c'_{t+1} | s_{t+1}) \right. \\ \left. \prod_i P(a'_{t+1} | s_{t+1}, c_{t+1}) \bar{P}(s_{t+2} | a_{t+1}, c_{t+1}), \forall (x_{t+1}, s_{t+2}) \in \mathcal{X} \times \mathcal{S} \right\} \quad (10)$$

where  $\mathcal{P}^\circ$  is the interior of  $\mathcal{P}$  and  $\bar{P}$  indicates that this distribution is fixed.

In Langer and Ay (2021), we iteratively project between these two sets to find the distribution in  $\mathcal{M}_{\mathcal{A}}^P$  that is closest to  $\mathcal{M}_G^P$ . This would be the distribution that describes a valid agent and has a high likelihood of achieving the goal. This approach is also called *planning as inference* (Attias, 2003; Toussaint, 2009; Toussaint et al., 2006). The approach is guaranteed to converge but might converge to a local minimum.

In our case, we want to adapt this approach to simultaneously learn the internal world model. The distribution  $P(S'_{t+1} | A_t, C_t)$  predicts the next sensory input and reflects, therefore, the agent's understanding of its environment. Hence we want to optimize our world model such that

$$P(S'_{t+2} | S_{t+1}, A_{t+1}) = \tilde{P}(S_{t+2} | S_{t+1}, A_{t+1}) \quad (11)$$

where  $\tilde{P}$  is the sampled ideal world model. We sample the distributions  $P(S_t, A_t, C_t)$  and  $P(S_{t+1} | S_t, A_t)$  as described in more detail by Zahedi et al. (2010), and we mark sampled distributions by a tilde,  $\tilde{P}$ . Note that we require the goal to be defined by a joint distribution, not a conditional, hence the actual optimization works with

$$\bar{P}(S_{t+1}, A_{t+1}) \otimes P(S'_{t+2} | S_{t+1}, A_{t+1}) = \bar{P}(S_{t+1}, A_{t+1}) \otimes \tilde{P}(S_{t+2} | S_{t+1}, A_{t+1}) \quad (12)$$

The joint distribution  $\bar{P}(S_{t+1}, A_{t+1})$  is fixed to the joint distribution resulting from the last step in the algorithm.

The conditional distribution  $P(S'_{t+2} | S_{t+1}, A_{t+1})$  can be calculated as

$$P(S_{t+2} | S_{t+1}, a_{t+1}) = \frac{\sum_{c_{t+1}} P_{a_t, c_t}(s_{t+1}) P_{c_t}(c_{t+1} | s_{t+1}) P(a_{t+1} | s_{t+1}, c_{t+1}) P(s_{t+2} | c_{t+1}, a_{t+1})}{\sum_{c_{t+1}} P_{a_t, c_t}(s_{t+1}) P_{c_t}(c_{t+1} | s_{t+1}) P(a_{t+1} | s_{t+1}, c_{t+1})} \quad (13)$$

Then the world goal manifold results in

$$\mathcal{M}_G^W(x_t, \bar{P}) = \{Q \in \mathcal{P}(\mathcal{X} \times \mathcal{S}) \mid Q(s_{t+2}, s_{t+1}, a_{t+1}) = \bar{P}(s_{t+1}, a_{t+1}) \tilde{P}(s_{t+2} | s_{t+1}, a_{t+1}), \\ \forall (s_{t+1}, a_{t+1}, s_{t+2}) \in \mathcal{S} \times \mathcal{A} \times \mathcal{S}\} \quad (14)$$

Similar to the agent manifold, we also define a world agent manifold:

$$\mathcal{M}_{\mathcal{A}}^W(x_t, \bar{P}) = \left\{ P \in \mathcal{P}^\circ(\mathcal{X} \times \mathcal{S}) \mid P(x_{t+1}, s_{t+2}) = P_{a_t, c_t}(s_{t+1}) \prod_j \bar{P}_{c_t}(c_{t+1}^j | s_{t+1}) \right. \\ \left. \prod_i \bar{P}(a_{t+1}^i | s_{t+1}, c_{t+1}) P(s_{t+2} | a_{t+1}, c_{t+1}), \forall (x_{t+1}, s_{t+2}) \in \mathcal{X} \times \mathcal{S} \right\} \quad (15)$$

Note that we can define a full agent manifold by

$$\mathcal{M}_{\mathcal{A}}(x_t) = \left\{ P \in \mathcal{P}^\circ(\mathcal{X} \times \mathcal{S}) \mid P(x_{t+1}, s_{t+2}) = P_{a_t, c_t}(s_{t+1}) \prod_j P_{c_t}(c_{t+1}^j | s_{t+1}) \right. \\ \left. \prod_i P(a_{t+1}^i | s_{t+1}, c_{t+1}) P(s_{t+2} | a_{t+1}, c_{t+1}), \forall (x_{t+1}, s_{t+2}) \in \mathcal{X} \times \mathcal{S} \right\} \quad (16)$$

and then  $\mathcal{M}_{\mathcal{A}}^W \subset \mathcal{M}_{\mathcal{A}}$  and  $\mathcal{M}_{\mathcal{A}}^G \subset \mathcal{M}_{\mathcal{A}}$ . Similarly, we can also define a full world goal manifold:

$$\mathcal{M}_{\mathcal{G}}^W(x_t) = \left\{ Q \in \mathcal{P}(\mathcal{X} \times \mathcal{S}) \mid \exists P \in \mathcal{P}(\mathcal{X}) : Q(s_{t+2}, s_{t+1}, a_{t+1}) = P(s_{t+1}, a_{t+1}) \right. \\ \left. \tilde{P}(s_{t+2} | s_{t+1}, a_{t+1}) \forall (s_{t+1}, a_{t+1}, s_{t+2}) \in \mathcal{S} \times \mathcal{A} \times \mathcal{S} \right\} \quad (17)$$

Now we are able to define the algorithm depicted in Figure A1.

The first step is to project to  $\mathcal{M}_{\mathcal{G}}^P(x_t)$  via an  $\epsilon$ -projection:

$$Q^0 = \arg \inf_{Q \in \mathcal{M}_{\mathcal{G}}^P(x_t)} D(Q \| P^0) \quad (18)$$

Then we project with an  $m$ -projection to  $\mathcal{M}_{\mathcal{A}}^P(x_t, P^0)$ :

$$P^1 = \arg \inf_{P \in \mathcal{M}_{\mathcal{A}}^P(x_t, P^0)} D(Q^0 \| P) \quad (19)$$

Up to this point, this is the standard em-algorithm. Now, instead of projecting to  $\mathcal{M}_{\mathcal{G}}^P(x_t)$  again, we update the internal world model by projecting to  $\mathcal{M}_{\mathcal{G}}^W(x_t, P^1)$  with an  $\epsilon$ -projection:

$$Q^1 = \arg \inf_{Q \in \mathcal{M}_{\mathcal{G}}^W(x_t, P^1)} D(Q \| P^1) \quad (20)$$

Afterward, we project, with an  $m$ -projection to  $\mathcal{M}_{\mathcal{A}}^W(x_t, P^1)$ ,

$$P^2 = \arg \inf_{P \in \mathcal{M}_{\mathcal{A}}^W(x_t, P^1)} D(Q^1 \| P) \quad (21)$$

Then we start the process over by projecting to  $\mathcal{M}_{\mathcal{G}}^W(x_t)$ , as depicted in Figure A1.

Note that this algorithm is not guaranteed to converge. Because we are interested in agents that learn while performing a task, we perform only the five optimization steps, as described earlier, after each step of the agent. Therefore convergence is not needed in our scenario.

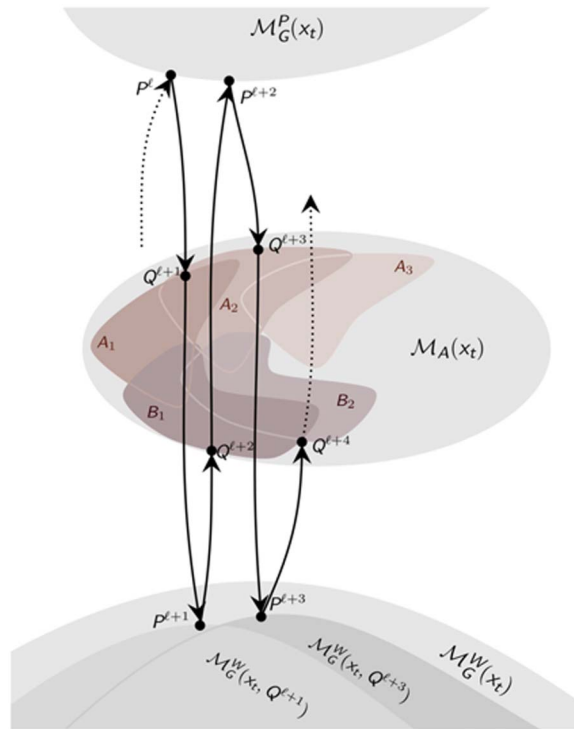


Figure A1. The modified em-algorithm for optimizing the behavior and the internal world model simultaneously. Here  $A_1 = \mathcal{M}_A^G(x_t, p^t)$ ,  $A_2 = \mathcal{M}_A^G(x_t, p^{t+2})$ ,  $A_3 = \mathcal{M}_A^G(x_t, p^{t+4})$ ,  $B_1 = \mathcal{M}_A^W(x_t, p^{t+1})$ , and  $B_2 = \mathcal{M}_A^W(x_t, p^{t+3})$ .