Chapter 9
On the Causal Structure of the Sensorimotor Loop

Nihat Ay and Keyan Zahedi

9.1 Introduction

In recent years, the application of information theory to the field of embodied intelligence has turned out to be extremely fruitful. Here, several measures of information flow through the sensorimotor loop of an agent are of particular interest. There are mainly two ways to apply information theory to the sensorimotor setting.

First, information-theoretic measures can be used within various analysis methods. Sensorimotor interactions of an embodied agent lead to the emergence of redundancy and structure of the agent's intrinsic processes. Understanding the generation of structure in the sensorimotor process and its exploitation is important within the field of embodied intelligence (Pfeifer and Bongard 2006). The quantification and analysis of information flows through an agent's sensorimotor loop from the perspective of an external observer, that is from the perspective of a scientist, proves to be effective in this regard (Lungarella and Sporns 2005, 2006). Here, transfer entropy (Schreiber 2000) has been used in order to quantify the flows of information between various processes of the sensorimotor loop, such as the sensor process on the actuator process. Furthermore, excess entropy, also known as predictive information (Bialek et al. 2001), has been used to analyse the interplay between information-theoretic measures and behavioral patterns of embodied agents (Der et al. 2008).

Second, information-theoretic measures can be used as objective functions for self-organized learning. This is based on the hypothesis that learning in natural intelligent systems is partly governed by an information-theoretic optimisation principle. Corresponding studies aim at the implementation of related principles, so-called

Nihat Ay · Keyan Zahedi
Max Planck Institute for Mathematics in the Sciences,
Inselstraße 22, D-04103, Leipzig, Germany
e-mail: {nay,zahedi}@mis.mpg.de

Nihat Ay
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA

M. Prokopenko (ed.), *Guided Self-Organization: Inception,*
Emergence, Complexity and Computation 9,
DOI: 10.1007/978-3-642-53734-9_9,

Infomax principles, in artificial systems. Emergent structures at various levels are then analysed in view of corresponding biological structures. In the sensorimotor setting, predictive information maximization has been used as a driving force for self-organised learning (Ay et al. 2008; Zahedi et al. 2010; Ay et al. 2012; Martius et al. 2013). As a result, the emergence of coordinated behavior with distributed control has been shown. The excess entropy has also been applied to similar systems within an evolutionary optimisation context (Prokopenko et al. 2006). Other measures of interest are the notion of relevant information (Polani et al. 2006) and empowerment (Klyubin et al. 2005). In the latter case, the maximization of empowerment determines the behavior of the agent and is not the basis of learning.

Most of the information-theoretic quantities mentioned above have the mutual information of two variables as an important building block. In information theory, this fundamental quantity is used as a measure of the transmission rate of a sender-receiver channel. Therefore, the objective functions for learning that are based on mutual information are usually associated with some kind of information flow. There is one problem with this interpretation, which is not visible in the simple sender-receiver context. Information flows are causal in nature, and related measures should be consistent with this fact. However, the causal aspects are usually not explicitly addressed. In order to do so, it has been proposed to combine information theory with the theory of causal networks (Ay and Polani 2008), based on the causal structure of the sensorimotor loop (Klyubin et al. 2004). This combination allows us to understand how stochastic dependence, and, in particular, the statistical structure of the sensorimotor process, is built up by causal relationships. Various (associational) information flow measures can be formulated in causal terms and lead, in general, to a modification of these measures. Thereby, the repertoire of information-theoretic quantities that can be used within the above-mentioned lines of research is extended. However, currently it is not clear to what extent realistic objective functions for self-organised learning should be causal in nature.

In Section 9.2 we sketch the theory of causal networks and its application to the sensorimotor loop setting. Section 9.3 introduces the notion of a causal effect and the identifiability problem of causal effects. In particular, the identifiability of causal effects from the intrinsic perspective of an agent is discussed. In Section 9.4 basic information-theoretic quantities are introduced. Within the context of temporal processes, such as the sensorimotor process, transfer entropy and predictive information are highlighted as important quantities. The maximization of predictive information is studied in an experimental setup in the final Section 9.5.

9.2 Causal Networks

9.2.1 The Definition of Causal Networks

The formal tool that we use for modelling causality is given in terms of Bayesian networks (Pearl 2000). They have two components, a structural and a functional one. The structural component is given in terms of a network. The network consists

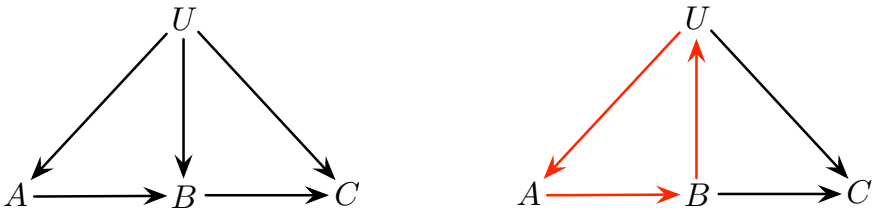


Fig. 9.1 Left: A network without a directed cycle, referred to as directed acyclic network (DAG). Right: A network with a directed cycle.

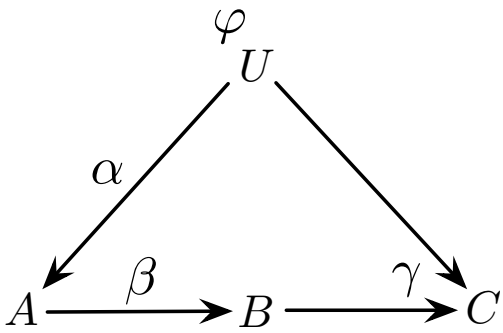


Fig. 9.2 Causal network

of vertices, or nodes, and directed edges. We denote the vertex set by V and the edge set by E which is formally a set of pairs $(v,w) \in V \times V$. Here, the pair (v,w) denotes the edge from v to w. The causal interpretation is that v is a *direct cause* of w and w is a *direct effect* v. The direct causes of a node v are referred to as *parents of* v and denoted by $pa(v)$. Extending this direct cause-effect relation, we say that a node v is a *cause* of a node w, and that w is an *effect* of v, if there is a directed path from v to w, denoted by $v \rightsquigarrow w$ (here, we exclude paths of length 0). According to this causal interpretation of the network we have to postulate that an effect cannot precede its cause. Stated differently, if v is a cause of w then w cannot be a cause of v. This is equivalent to the property that the network does not have any directed cycles as shown in Figure 9.1. A directed network with this property is called a *directed acyclic graph* (*DAG*).

Given a node v with state set $\mathscr{X}_v$, the mechanism of v is formalized in terms of a stochastic map $\kappa(x;x'), x \in \mathscr{X}_{pa(v)}, x' \in \mathscr{X}_v$, that is $\sum_{x'} \kappa(x;x') = 1$ for all x. We refer to these maps also as (*Markov*) *kernels*. Before we provide a general definition of a Bayesian network, we illustrate its basic concepts in terms of an instructive example. To this end, we consider the DAG with the nodes U, A, B, and C that is shown in Figure 9.2. In addition to the graph, the mechanisms φ, α, β, and γ are given. They describe how the nodes function and are formalized in terms of stochastic maps. For example, $\gamma(u,b;c)$ stands for the probability that node C is in state c given that it

has received b and u. Based on these mechanisms, the probability of observing the states u,a,b, and c in the unperturbed system can be computed as the product

$$p(u,a,b,c) \;=\; \varphi(u)\cdot\alpha(u;a)\cdot\beta(a;b)\cdot\gamma(u,b;c)\,. \tag{9.1}$$

This equation connects the phenomenological level (left-hand side of equation (9.1)) and the mechanistic level (the individual terms on the right-hand side of equation (9.1)).

Now we come to the general setting of a Bayesian network. Bayesian networks are based on DAGs but have a further structure as model of the involved mechanisms. As in the previous example of Figure 9.2, in a Bayesian network to each node v a mechanism κ^v is assigned. For simplicity of the arguments and derivations, we assume that the nodes v have finitely many states $\mathscr{X}_v$. Each node gets inputs $x_{pa(v)}$ from the parents $pa(v)=\{u\in V:(u,v)\in E\}$ and generates a stochastic output according to the distribution $\kappa^v(x_{pa(v)};\cdot)$. All the mechanisms together generate a distribution of global states. In order to describe this distribution, we choose a numbering of the nodes, that is $v_1,v_2,\ldots,v_n$, which is compatibe with the causal order given by the graph. More precisely, we assume the following: if there is a directed path from v_i to v_j then i is smaller than j. We use this numbering in order to generate the states of the individual nodes.

$$p(x_{v_1},x_{v_2},\ldots,x_{v_n}) \;=\; \kappa^{v_1}(x_{v_1})\cdot\kappa^{v_2}(x_{pa(v_2)};x_{v_2})\cdots\kappa^{v_n}(x_{pa(v_n)};x_{v_n}).$$

This is clearly independent of the particular choice of such an admissible numbering. Therefore, we can write

$$p(x_v : v\in V) \;=\; \prod_{v\in V}\kappa^v(x_{pa(v)};x_v). \tag{9.2}$$

On the left-hand side of this equation we have the probability of observing a particular global configuration x_v, $v\in V$. On the right-hand side we have the mechanisms. The equation postulates the transition from the mechanistic level to the phenomenological level. Given a joint distribution describing the phenomenological level, there is always a Bayesian network that generates that distribution. In order to see this, choose an arbitrary ordering $v_1,v_2,\ldots,v_n$ of the nodes. The following equality holds in any case (whenever the conditional probabilities on the right-hand side are defined):

$$p(x_{v_1},x_{v_2},\ldots,x_{v_n}) \;=\; \prod_{i=1}^{n} p(x_{v_i}\,|\,x_{v_1},\ldots,x_{v_{i-1}}). \tag{9.3}$$

Consider now the graph in which a pair (v_i,v_j) is an edge if and only if i is smaller than j. With respect to this graph, the parent set of a node v_i is given by $v_1,v_2,\ldots,v_{i-1}$. Defining stochastic maps κ^{v_i} with

$$\kappa^{v_i}(x_{pa(v_i)};x_{v_i}) \;:=\; p(x_{v_i}\,|\,x_{v_1},\ldots,x_{v_{i-1}}),$$

whenever $p(x_{v_1},\ldots,x_{v_{i-1}}) > 0$, the equation (9.3) reduces to (9.2). Note that there are many possible mechanistic explanations of a given joint distribution. Having a particular one only means that *one possible* explanation is given, which does not necessarily represent *the* actual mechanisms that underly the joint distribution.

Definition 1 (Causal Markov Property). Given a DAG $G = (V,E)$, we say that a probability measure p on $\times_{v\in V}\mathscr{X}_v$ satisfies the *causal Markov property*, if, with respect to p, each variable is stochastically independent of its non-effects (V minus set of effects), conditional on its direct causes.

This property is also referred to as *local Markov property*. The conditional independence statements of the local Markov property imply also other conditional independence statements that can be deduced from the graph. In order to be more precise, we have to introduce the notion of *d-separation*.

Definition 2. Let $G = (V,E)$ be a DAG, and let S be a (possibly empty) subset of V. We say that a path $(v_1,\ldots,v_k)$ is *blocked* by S if there is a node v_i on the path such that

- either $v_i \in S$, and edges of the path do not meet head-to-head at v_i, or
- v_i and all its descendants are not in S, and edges of the path meet head-to-head at v_i.

Two non-empty and disjoint sets $A,B \subseteq V \setminus S$ are *d-separated* by S if all paths between A and B are blocked by S.

The notion of d-separation is completely structural. The local Markov property provides a way of coupling the structure with the joint probability distribution p. This condition implies a seemingly stronger Markov property: we say that p satisfies the *global Markov property*, if

A and B are d-separated by $S \Longrightarrow A$ and B are stochastically independent given S.

The following theorem is central in graphical models theory (Lauritzen 1996).

Theorem 1. *Let $G = (V,E)$ be a DAG. For a probability measure p on $\times_{v\in V}\mathscr{X}_v$, the following conditions are equivalent:*

1. *p admits a factorization according to G (a factorization like in formula (9.2)).*
2. *p obeys the global Markov property, relative to G.*
3. *p obeys the causal Markov property, relative to G.*

A simple application of this theorem to the causal graph of Figure 9.2 yields the following conditional independence statements:

U and B are stochastically independent given A. (9.4)

A and C are stochastically independent given U and B. (9.5)

These and similar conditional independence statements will be used in the context of the sensorimotor loop.

9.2.2 The Causal Structure of the Sensorimotor Loop

The following figure (Figure 9.3) illustrates the components of a sensorimotor loop with their respective interactions. In order to apply the theory of causal networks to the sensorimotor loop, we have to consider a causal network that captures the main aspects of this structure. Figure 9.4 shows the general causal network of a sensorimotor loop, where W_t, S_t, C_t, A_t denote the state of the world, the sensor, the controller, and the actuator at some time point t, respectively (Klyubin et al. 2004; Ay and Polani 2008). We denote the corresponding state spaces by $\mathscr{W}, \mathscr{S}, \mathscr{C}$, and $\mathscr{A}$. The stochastic maps α, β, φ, and π describe the mechanisms that are involved in the sensorimotor dynamics:

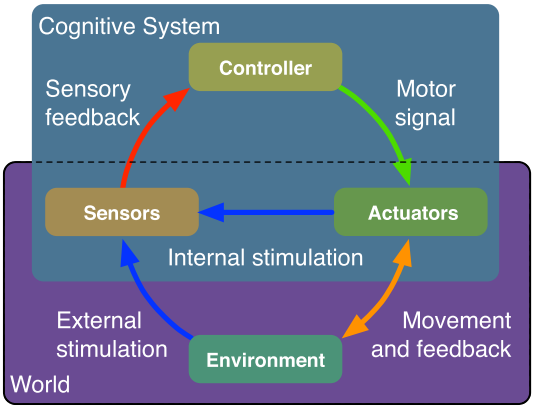


Fig. 9.3 Sensorimotor loop

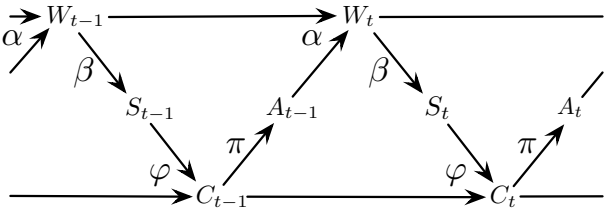


Fig. 9.4 Sensorimotor loop

$$\alpha : \mathscr{W} \times \mathscr{A} \to \mathcal{P}(\mathscr{W}),$$
$$\beta : \mathscr{W} \to \mathcal{P}(\mathscr{S}),$$
$$\varphi : \mathscr{C} \times \mathscr{S} \to \mathcal{P}(\mathscr{C}),$$
$$\pi : \mathscr{C} \to \mathcal{P}(\mathscr{A}).$$

Here, $\mathcal{P}(\mathscr{X})$ denotes the set of probability measures on $\mathscr{X}$. The kernels α and β encode the constraints of the sensorimotor loop due to the agent's morphology and the properties of its environment. The mechanisms α and β are extrinsic and encode the agent's embodiment which sets constraints for the agent's behavior and learning. The kernels φ, π are intrinsic with respect to the agent and are assumed to be modifiable through a learning process (for details, see (Zahedi et al. 2010)).

The process (W_t, S_t, C_t, A_t), $t = 0, 1, 2, \ldots$, is generated by Markov transition kernels. Given the above kernels as models of the mechanisms that constitute the sensorimotor loop, and given an initial distribution μ, we obtain the joint distribution

$$p(w_0, s_0, c_0, a_0, \ldots, w_n, s_n, c_n, a_n)$$
$$= \mu(w_0, s_0, c_0, a_0) \cdot \prod_{t=1}^{n} \alpha(w_{t-1}, a_{t-1}; w_t)\, \beta(w_t; s_t)\, \varphi(c_{t-1}, s_t; c_t)\, \pi(m_t; a_t).$$

As examples, we now consider sensorimotor loops where particular arrows are removed.

Example 1. 1. *Passive observer.* If the agent does not act on the world at all or, stated differently, if it only observes its environment then there is no edge from A_t to W_t and the kernel α does not involve the actuator state. More precisely,

$$\alpha(w, a; w') = \alpha(w; w') \quad \text{for all } w, a, \text{ and } w'.$$

Obviously, in this situation we can remove the edges from A to W and from C to A, which leads to the diagram shown in Figure 9.5.

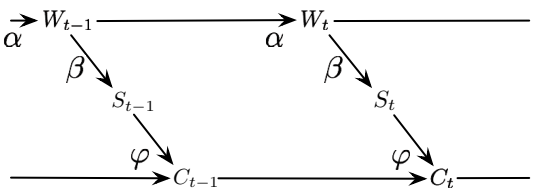


Fig. 9.5 Sensorimotor loop of a passive observer

2. *Open loop controller.* The situation in which the agent does not sense anything in the world is referred to as *open loop control*. Here, the kernel φ does not use the sensor state s, that is

$$\varphi(c,s;c') = \varphi(c;c') \quad \text{for all } c,\ s, \text{ and } c'.$$

Here, the edges from S to C and from W to S can be removed, and we obtain the diagram shown in Figure 9.6.

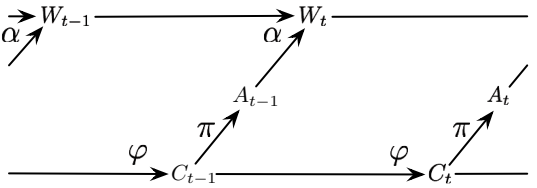


Fig. 9.6 Sensorimotor loop of an open loop controller

3. *Memoryless or reactive controller.* In this example, we assume that there is no edge from the controller state at time $t-1$ to the controller state at time t (see Figure 9.7 (A)). This means

$$\varphi(c,s;c') = \varphi(s;c') \quad \text{for all } c,\ s, \text{ and } c'.$$

If we combine the kernels φ and π to one kernel, which we again denote by π, then we have a representation of a memoryless controller that is often referred to as *reactive controller* (see Figure 9.7 (B)).

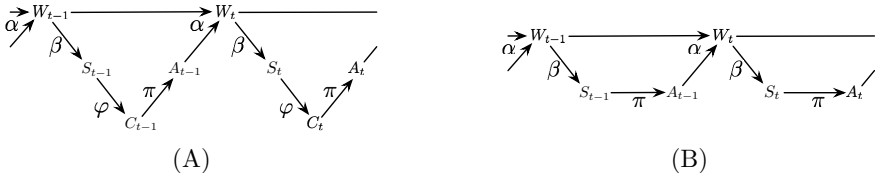


Fig. 9.7 Sensorimotor loop of a memoryless or reactive controller

9.3 Causal Effects

9.3.1 The Definition of Causal Effects

In order to study causal effects, one has to apply an interventional operation, which we also call *clamping*. Clamping the state of a node means a change of the mechanism of that node, and it is formalized by the so-called *do*-operation. In order to explain the main idea behind the *do*-operation, we first consider our example of Figure 9.2. In this example we compute the causal effect of A on C. If we clamp $\hat{a}$, or,

in Pearl's terminology (Pearl 2000), do $\hat{a}$, we would have to replace equation (9.1) by

$$p_{\hat{a}}(u,a,b,c) \;=\; \varphi(u)\cdot\hat{\alpha}(a)\cdot\beta(a;b)\cdot\gamma(u,b;c)\,, \tag{9.6}$$

where $\hat{\alpha}(a)=1$, if $a=\hat{a}$, and $\hat{\alpha}(a)=0$, if $a\neq\hat{a}$. In particular, after clamping node A the new mechanism $\hat{\alpha}$ is not sensitive to u anymore. In terms of the *do*-formalism, the post-interventional probability measure $p_{\hat{a}}(u,a,b,c)$ is written as $p_{\hat{a}}(u,a,b,c) = p(u,a,b,c\,|\,do(\hat{a}))$. Summation over u,a,b yields the probability of observing c after having clamped $\hat{a}$:

$$p(c\,|\,do(\hat{a})) \;=\; \sum_{u,b}\varphi(u)\cdot\beta(\hat{a};b)\cdot\gamma(u,b;c)\,. \tag{9.7}$$

We refer to this post-intervational probability measure as *causal effect* of A on C (see Figure 9.8).

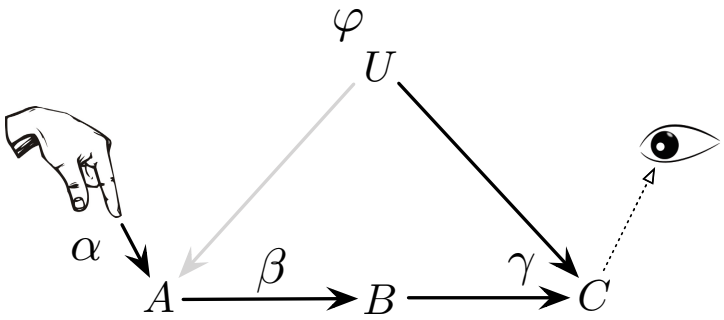


Fig. 9.8 Intervention

Now, we extend this idea of intervention to the general setting of Bayesian networks. Knowing the mechanisms, it is possible to model the consequences of intervention. Consider the equation (9.2), a subset A of V, and assume that the configuration of A is externally set to x_A. What does this mean? It means that the mechanisms of the nodes v in A are replaced by the mechanisms

$$\delta_{x_v}(x'_v) \;=\; \begin{cases} 1, & \text{if } x'_v = x_v \\ 0, & \text{otherwise} \end{cases}.$$

This leads to the truncated product

$$p(x_{V\setminus A}\,|\,do(x_A)) \;=\; \prod_{v\in V\setminus A}\kappa^v(x_{pa(v)};x_v)\,. \tag{9.8}$$

Although it is not essential for what follows, we present the formula for the usual way of conditioning:

$$p(x_{V\setminus A} \mid x_A) = \frac{p(x_{V\setminus A}, x_A)}{p(x_A)} = \frac{\prod_{v\in V} \kappa^v(x_{pa(v)}; x_v)}{\sum_{x'_V : x'_{V\setminus A} = x_{V\setminus A}} \prod_{v\in V} \kappa^v(x'_{pa(v)}; x'_v)}. \tag{9.9}$$

This shows that the interventional conditioning, given by equation (9.8), is much easier to compute than the standard conditioning of equation (9.9).

In the above derivations we considered a set A where the intervention takes place and observed the complement $V \setminus A$ of A, which corresponds to the causal effect $p(x_{V\setminus A} \,|\, do(x_A))$ (see Figure 9.9 (A)). In more general situations, the post-interventional distribution is observed in a subset B of $V \setminus A$. In order to define this, we simply have to marginalize out the unobserved nodes:

$$p(x_B \,|\, do(x_A)) = \sum_{x_{V\setminus(A\cup B)}} p(x_B, x_{V\setminus(A\cup B)} \,|\, do(x_A)). \tag{9.10}$$

Comparing this interventional conditioning with classical conditioning, we observe one important difference. The latter is only possible when the event that is conditioned on has positive probability. It describes the probability of observing an event F if an event E has already been observed. However, if the probability for E vanishes then, already at an intuitive level, it is not clear with which probability the occurrence of F should be expected. Formally, if $p(x_A) = 0$ then the expression (9.9) is not well defined. The interventional conditioning is different. Already at an

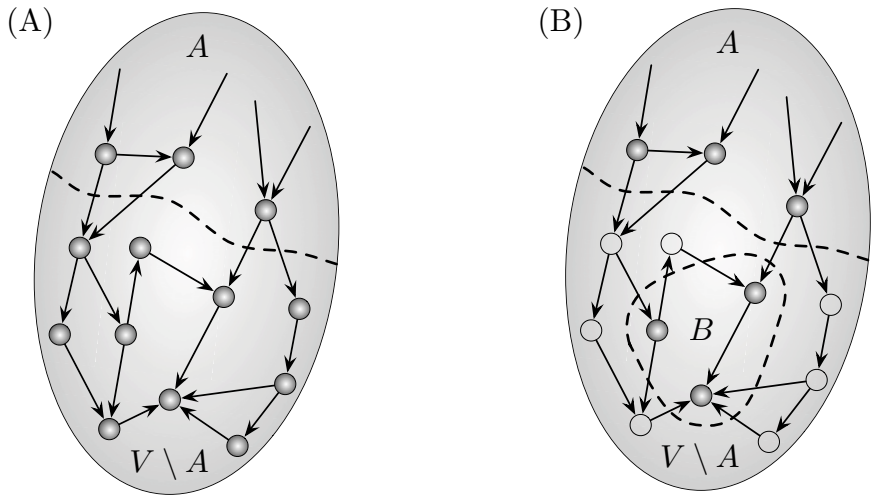


Fig. 9.9 (A) intervention in A and observation in the complement $V \setminus A$ of A in V, (B) intervention in A and observation in B

intuitive level it is clear that any intervention will lead to some reaction of the system. Formally, we see this in equation (9.8). The product on the right-hand side of this equation is always well defined.

Although interventional conditioning differs from observational conditioning, in some cases both operations coincide. In order to be more precise, we note that any observed association of two variables A and B in a Bayesian network has basically three sources: A is a cause of B, B is a cause of A, or there is a common cause of A and B. This is known as the common cause principle (Reichenbach 1956). If we assume that B is not a cause of A and there is no common cause of A and B then there is no explanation for the association of A and B other than A being a cause of B. In that case, all stochastic dependence between A and B is due to the causal effect of A on B. For such a situation we say that the pair (A,B) is a *causal pair*. One can show that for causal pairs interventional and observational conditioning are equivalent.

Proposition 1. *If (A,B) is a causal pair with respect to a DAG G, then for any Bayesian network with graph G, we have $p(x_B \,|\, do(x_A)) = p(x_B \,|\, x_A)$.*

This proposition shows that in some cases the post-interventional distribution, that is $p(x_B \,|\, do(x_A))$, can be obtained by observation only. More precisely, if observations of all joint events (x_A,x_B) allow us to estimate their probabilities $p(x_A,x_B)$, then one can compute the causal effect as

$$p(x_B \,|\, do(x_A)) = \frac{p(x_A,x_B)}{\sum_{x'_B} p(x_A,x'_B)}, \tag{9.11}$$

whenever $p(x_A) = \sum_{x'_B} p(x_A,x'_B) > 0$. If a causal effect can be computed in this way, that is without intervention, we say that it is identifiable. In the next section, this subject is discussed to the extent to which it is used in the context of the sensorimotor loop. We will argue that causal effects that are relevant to an agent should be identifiable with respect to the intrinsic variables of that agent.

9.3.2 Identification of Causal Effects

The equation (9.8) provides a formal definition of a causal effect. Such a causal effect can be determined in various ways depending on the available experimental operations. If we can experimentally intervene into the system, then the mechanisms will generate the post-interventional probability measure which can be observed. In many cases, however, experimental intervention is not possible. Then one has to ask the following question: *Is it possible to conclude the consequences of intervention solely on the basis of observation, that is without actual intervention?* At first sight, the answer seems to be clearly *No!* In some sense, this *is* already the whole answer to this question. On the other hand, Proposition 1 proves that in the case of a causal pair, it is indeed possible to compute the causal effect, left-hand side of (9.11), without intervention, from the right-hand side of (9.11). This demonstrates that, in order to identify a causal effect observationally, we require some structural information.

Without any structural information, it is not possible at all to identify causal effects. However, this does not mean that we need to know the complete structure, that is the DAG, in order to identify a causal effect. Partial structural knowledge can be sufficient for the identification of a causal effect, as we will see.

In what follows we further illustrate the problem of causal effect identification using our standard example. We want to compute the causal effect of A on C. Obviously, U is a common cause of A and C, and therefore the pair (A,C) is not a causal pair, which implies $p(c\,|\,do(a)) \neq p(c\,|\,a)$ in general. The following formula shows a different way of obtaining $p(c\,|\,do(a))$ by assuming that we can observe A, B, and C:

$$p(c\,|\,do(a)) \;=\; \sum_b p(b\,|\,a) \sum_{a'} p(a') \cdot p(c\,|\,a',b). \tag{9.12}$$

This equality is quite surprising and not obvious at all. Here is the proof:

$$\begin{aligned}
p(c\,|\,do(a)) &= \sum_{u,b} \varphi(u) \cdot \beta(a;b) \cdot \gamma(u,b;c) \qquad \text{(formula (9.7))} \\
&= \sum_{u,b} p(u) \cdot p(b\,|\,a) \cdot p(c\,|\,u,b) \\
&= \sum_{u,b} \left(\sum_{a'} p(a') \cdot p(u\,|\,a') \right) p(b\,|\,a) \cdot p(c\,|\,u,b) \\
&= \sum_b p(b\,|\,a) \sum_{a'} p(a') \sum_u p(u\,|\,a') \cdot p(c\,|\,u,b) \\
&= \sum_b p(b\,|\,a) \sum_{a'} p(a') \sum_u p(u\,|\,a',b) \cdot p(c\,|\,u,b,a') \\
&\quad \text{(conditional inedependence statements (9.4) and (9.5))} \\
&= \sum_b p(b\,|\,a) \sum_{a'} p(a') \cdot p(c\,|\,a',b)
\end{aligned}$$

The structure of this example will be revisited in the context of the sensorimotor loop.

9.3.3 Causal Effects in the Sensorimotor Loop

In this section we study the problem of the identification of causal effects in the sensorimotor loop. In this context, there are various causal effects of interest, for instance the effect of actions on sensor inputs. We are mainly interested in causal effects that involve intrinsic variables of the agent, that is the variables S_t, C_t, and A_t, $t = 0,1,2,\ldots$. Other causal effects can not be evaluated by the agent, because extrinsic variables are by definition not directly available to the agent. In the proposition below, we list three causal effects in the sensorimotor loop that are identifiable by the agent without actual intervention and purely based on in situ observations of the agent. In order to be more precise, we have a closer look at the causal diagram of the transition from time $t-1$ to t. Here, as shown in Figure 9.10, we consider the future sensor value of only one time step and summarize the past process by one

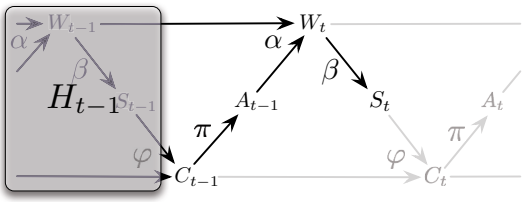


Fig. 9.10 Reduction procedure of the causal diagram

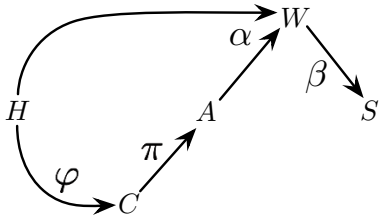


Fig. 9.11 Reduced causal diagram for one time step

variable H_{t-1}. We focus on the resulting causal diagram of Figure 9.11. The joint distribution in the reduced diagram is given as

$$p(h,c,a,w,s) \;=\; p(h)\,\varphi(h;c)\,\pi(c;a)\,\alpha(h,a;w)\,\beta(w;s). \tag{9.13}$$

We can now consider, for instance, the causal effect of A on S. In general, we do *not* have $p(s\,|\,do(a)) = p(s\,|\,a)$, which follows from the fact that H is a common cause of A and S. Nevertheless, it turns out that this causal effect, among others, is identifiable with respect to the intrinsic variables of the agent.

Proposition 2. *Let the joint distribution (9.13) be strictly positive. Then the following equations hold:*

$$\textbf{(1)} \qquad p(s\,|\,do(a),c) := \frac{p(s,c\,|\,do(a))}{p(c\,|\,do(a))} \;=\; p(s\,|\,c,a)$$

$$\textbf{(2)} \qquad p(s\,|\,do(a)) \;=\; \sum_c p(s\,|\,c,a)\,p(c)$$

$$\textbf{(3)} \qquad p(s\,|\,do(c)) \;=\; \sum_a p(a\,|\,c) \sum_{c'} p(s\,|\,c',a)\,p(c').$$

The proof of Proposition 2 is given in the appendix. In all three causal effects of this proposition, the conditional distribution $p(s\,|\,c,a)$ turns out to be essential as building block for the identification of the causal effects. It describes the expectation of the agent to observe s, given that it is in state c and performs an action a. We refer to this conditional probability distribution as *world model* of the agent. Note that in the strictly positive case, according to Proposition 2 (1), the world model is not

dependent on the agent's policy. These results indicate that the world model plays an important role in evaluating causal effects in the sensorimotor loop. Furthermore, it is an essential object within the empowerment approach to behavior (Klyubin et al. 2005). We will see that the world model also plays a fundamental role within learning processes.

9.4 Information Flows

9.4.1 Information-Theoretic Preliminaries

In this section, we introduce a few fundamental quantities known from information theory. First, we consider Shannon information. Assume that $p(x)$ describes the expectation that the outcome of a random experiment is going to be x. The information that we receive by knowing the outcome x is equal to the surprise about that outcome, which is quantified by $-\ln p(x)$. Events x that we expect to occur with very low probability will highly surprise us when they do occur. This is expressed by a large value of $-\ln p(x)$ for a low probability $p(x)$. The expectation value of this function is known as *Shannon entropy* or *Shannon information*:

$$H_p(X) := -\sum_{x\in\mathcal{X}} p(x)\ln p(x)\,.$$

The Shannon entropy quantifies the uncertainty about the outcome of the random variable X. Note that for $p(x)=0$, the term $p(x)\ln p(x)$ is not directly defined. However, the function $t\ln t$ can be continuously extended in $t=0$ with value equal to 0, which justifies the convention $0\ln 0 = 0$. If we know the outcome of a random experiment beforehand, that is if our expectation p is concentrated around one value x_0, then the Shannon entropy is small. On the other hand, if the distribution is equally spread over all events, then the uncertainty about the outcome of X is maximal. With the uniform distribution u and the cardinality n of $\mathcal{X}$, we have

$$H_u(X) = -\sum_{x\in\mathcal{X}} \frac{1}{n}\ln\frac{1}{n} = -\ln\frac{1}{n} = \ln n\,.$$

Now we consider two variables X and Y. With the expectation $p(x,y)$ we associate the measure $H_p(X,Y)$ of uncertainty. Now assume that we have observed the variable X and have a remaining uncertainty about Y given X. This is given as

$$H(Y|X) = -\sum_{x\in\mathcal{X}} p(x)\sum_{y\in\mathcal{Y}} p(y|x)\ln p(y|x)\,.$$

Observation of X reduces the uncertainty about the outcome of Y and we define the reduction of uncertainty by

$$I(X;Y) := H(Y) - H(Y|X)\,.$$

This quantity is symmetric and has various other representations:

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= \sum_{x \in \mathscr{X}} p(x) \sum_{y \in \mathscr{Y}} p(y|x) \ln p(y|x) - \sum_{y \in \mathscr{Y}} p(y) \ln p(y) \\ &= I(Y;X)\,. \end{aligned}$$

This quantity is referred to as *mutual information* and quantifies the stochastic dependence of variables X and Y. We have $I(X;Y) = 0$ if and only if X and Y are stochastically independent, that is $p(x,y) = p(x)\,p(y)$. Introducing a further variable Z, we can consider the conditional mutual information

$$I(Z;Y|X) \;=\; H(Z|X) - H(Z|X,Y)\,.$$

This vanishes if and only if Z and Y are conditionally independent given X.

The mutual information of two variables X and Y can be extended to more than two variables. Given variables X_v, $v \in V := \{1,2,\ldots,N\}$, we define the *multi-information* as

$$I(X_1;X_2;\ldots;X_N) \;:=\; \sum_{v=1}^{N} H(X_v) - H(X_1,\ldots,X_N).$$

This definition is clearly independent of the order of the X_v.

Information-theoretic quantities can be used to characterize conditional independence. They can also be used to quantify the deviation from (conditional) independence as a measure for (conditional) stochastic dependence. Combined with the notion of intervention, this also leads to measures of causal information flows (Ay and Polani 2008).

9.4.2 Transfer Entropy and Causality

In this section, we draw a close connection between the multi-information and another information-theoretic quantity which addresses causal aspects of interacting processes. This quantity has been studied as *transfer entropy* by Schreiber (Schreiber 2000) and a slightly different version of it has been called *directed information* by Massey (Massey 1990). Recently, a thermodynamic interpretation of transfer entropy has been provided (Prokopenko et al. 2013). Using a somewhat implicit terminology, we argue that a more careful consideration of causality is necessary for understanding the sources of stochastic dependence.

In order to simplify the arguments we first consider only a pair X_t, Y_t, $t = 1,2,\ldots$, of stochastic processes which we also denote by X and Y. The extension to more than two processes will be straight-forward. Furthermore, we use the notation X^t for the random vector $(X_1,\ldots,X_t)$ and similarly x^t for the particular outcome $(x_1,\ldots,x_t)$.

Given a time n, consider the following quantity, which we refer to as *mutual information rate*:

$$\begin{aligned}
\mathfrak{I}(X^n;Y^n) &:= \frac{1}{n} I(X^n;Y^n) && (9.14)\\
&= \frac{1}{n}\Big(H(X^n)+H(Y^n)-H(X^n,Y^n)\Big) && (9.15)\\
&= \frac{1}{n}\sum_{t=1}^{n}\Big\{H(X_t\,|\,X^{t-1})+H(Y_t\,|\,Y^{t-1})-H(X_t,Y_t\,|\,X^{t-1},Y^{t-1})\Big\} && (9.16)\\
&= \frac{1}{n}\sum_{t=1}^{n}\left\{\underbrace{I(X_t;Y_t\,|\,X^{t-1},Y^{t-1})}_{I}+\underbrace{T(Y^{t-1}\to X_t)}_{II}+\underbrace{T(X^{t-1}\to Y_t)}_{III}\right\} && (9.17)
\end{aligned}$$

with the *transfer entropy* terms (Schreiber 2000)

$$\begin{aligned}
T(Y^{t-1}\to X_t) &:= I(X_t;Y^{t-1}\,|\,X^{t-1})\\
T(X^{t-1}\to Y_t) &:= I(Y_t;X^{t-1}\,|\,Y^{t-1}).
\end{aligned}$$

Furthermore, using standard arguments we easily see that in the case of stationary processes the following limit exists:

$$\mathfrak{I}(X;Y) := \lim_{n\to\infty}\mathfrak{I}(X^n;Y^n).$$

Note that in the case of independent and identically distributed variables (X_t,Y_t), $t=1,2,\ldots$, the transfer entropy terms *II* and *III* in (9.17) vanish and the only contributions to $\mathfrak{I}(X^n;Y^n)$ then are the mutual informations $I(X_t;Y_t)$. In the case of stationary processes these mutual informations coincide and we have $\mathfrak{I}(X;Y)=I(X_1;Y_1)$. In this sense the quantity $\mathfrak{I}$ extends the mutual information of two variables to a corresponding measure for two processes.

In general, the first term I of (9.17) quantifies the stochastic dependence of X_t and Y_t after "screening off" the causes of X_t and Y_t that are intrinsic to the system, namely X^{t-1} and Y^{t-1}. Assuming the principle of common cause (Reichenbach 1956), which postulates that *all* stochastic dependences are based on causal interactions, we can infer common causes of X_t and Y_t that act from outside on the system, if the first term I is positive. If, on the other hand, the system is closed in the sense that

$$p(x_t,y_t\,|\,x^{t-1},y^{t-1}) \;=\; p(x_t\,|\,x^{t-1},y^{t-1})\,p(x_t\,|\,x^{t-1},y^{t-1}), \qquad t=1,2,\ldots,$$

then the first term in (9.17) vanishes and the transfer entropies *II* and *III* are the only contributions to $\mathfrak{I}(X^n;Y^n)$. They refer to causal interactions within the system. As example we consider the term *II*:

$$T(Y^{t-1}\to X_t) \;=\; H(X_t\,|\,X^{t-1})-H(X_t\,|\,X^{t-1},Y^{t-1}).$$

It quantifies the reduction of uncertainty about the outcome x_t if, in addition to the knowledge about the previous outcomes $x_1,\dots,x_{t-1}$ of X, also the previous outcomes $y_1,\dots,y_{t-1}$ of Y are known. Therefore, the transfer entropy $T(Y^{t-1}\to X_t)$ has been used as a measure for the causal effect of Y^{t-1} on X_t (Schreiber 2000), (Kaiser and Schreiber 2002), which is closely related to the concept of *Granger causality*. On the other hand, a direct interpretation of transfer entropy as a measure for causal effects has some shortcomings which are already mentioned in (Kaiser and Schreiber 2002) and further addressed in (Ay and Polani 2008) within the context of Pearl's (Pearl 2000) causality theory. Below, we will illustrate these shortcomings demonstrating the need for an alternative formalization of causality.

The definitions of the multi-information rate and the transfer entropy naturally generalize to more than two processes.

Definition 3. Let X_v be stochastic processes with state space $\mathscr{X}_v$, $v\in V$. We define the *transfer entropy* and the *multi-information rate* in the following way:

$$\begin{aligned} T(X_{V\setminus v}{}^{t-1}\to X_{v,t}) &:= I(X_{v,t};X_{V\setminus v}{}^{t-1}\,|\,X_v{}^{t-1}) \\ \mathfrak{I}(X_v{}^n : v\in V) &:= \frac{1}{n}\left(\sum_{v\in V} H(X_v{}^n) - H(X_V{}^n)\right) \\ &= \frac{1}{n}\sum_{t=1}^{n}\left\{ I(X_{v,t} : v\in V\,|\,X_V{}^{t-1}) + \sum_{v\in V} T(X_{V\setminus v}{}^{t-1}\to X_{v,t})\right\} \end{aligned}$$

and, if stationarity is assumed,

$$\mathfrak{I}(X_v : v\in V) := \lim_{n\to\infty}\mathfrak{I}(X_v{}^n : v\in V).$$

In order to illustrate the problem of interpreting transfer entropy in a causal way, consider a finite node set V and a set $E\subseteq V\times V$ of directed edges. This is a way of encoding the structure of the nodes' causal interactions. The mechanisms are described in terms of Markov transition kernels

$$\kappa^v : \mathscr{X}_{pa(v)}\times\mathscr{X}_v\to[0,1],\qquad (x_{pa(v)},x'_v)\mapsto\kappa^v(x_{pa(v)};x'_v).$$

They define a "global kernel" as follows:

$$\kappa : \mathscr{X}_V\times\mathscr{X}_V\to[0,1],\qquad (x,x')\mapsto\prod_{v\in V}\kappa^v(x_{pa(v)};x'_v).$$

With a stationary distribution p_1 of κ, that is $\sum_{x_1} p_1(x_1)\,\kappa(x_1;x_2)=p_1(x_2)$, we consider the process $X_{V,t}=(X_{v,t})_{v\in V}$, $t=1,2,\dots$, that satisfies

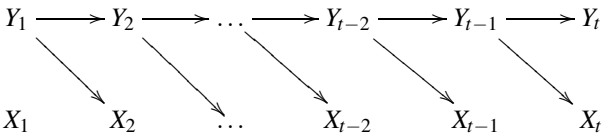
$$P\{X_1=x_1,\dots,X_n=x_n\} \;=\; p_1(x_1)\,\kappa(x_1;x_2)\cdots\kappa(x_{n-1};x_n).$$

In these definitions, the corresponding multi-information rate simplifies to

$$\mathfrak{I}(X_v{}^n : v \in V) = \frac{1}{n}\sum_{t=1}^{n}\sum_{v\in V} T(X_{V\setminus v}{}^{t-1} \to X_{v,t}).$$

If we assume that the individual kernels κ^v are deterministic then the global kernel κ is also deterministic. This implies that for all nodes v there is a time s such that all transfer entropies $T(X_{V\setminus v}{}^{t-1} \to X_{v,t})$, $t \geq s$, vanish, and thus the multi-information rate converges to zero. This appears counterintuitive because even if we have strong causal interactions within a deterministic system, the dynamics creates redundancy that allows all nodes v to predict their own next states $x_{v,t}$ from their previous states $x_{v,1}, \ldots, x_{v,t-1}$ without the need of additional information from the other nodes $V \setminus v$. This intrinsic prediction is not required to be mechanistically implemented and therefore screens off the actual mechanisms that might involve causal effects from the complement of v. We illustrate this effect by a more specific example.

We consider again two processes X and Y and assume, as illustrated in the following diagram, that the next state of X as well as the next state of Y only depend on the current state of Y.



In what follows we define a one-dimensional family of transition kernels. To this end, we first consider two extreme situations. In the first situation, we assume that there is no memory:

$$\kappa^X(x_{t-1}, y_{t-1}; x_t) \;=\; \kappa^Y(x_{t-1}, y_{t-1}; y_t) \;=\; \frac{1}{2}\,. \tag{9.18}$$

Obviously, in this situation there is no temporal information flow at all. The other extreme situation is given in the following way: In order to compute the next state, both nodes copy the current state of node Y and invert it.

$$(x, y) \to (-y, -y), \qquad x, y \in \{\pm 1\}. \tag{9.19}$$

In this case, the current state of Y completely determines the next state of X. Therefore, intuitively, one would expect a maximal amount of information flow from Y to X. We now interpolate these two extreme situations of minimal and maximal information flow in order to get a one-parameter family of transition kernels. More precisely, we define

$$\kappa^Y(x_{t-1}, y_{t-1}; y_t) \;:=\; \frac{1}{1+e^{2\beta y_t y_{t-1}}}, \qquad \kappa^X(x_{t-1}, y_{t-1}; x_t) \;:=\; \frac{1}{1+e^{2\beta x_t y_{t-1}}}.$$

Here, β plays the role of an inverse temperature. In the high-temperature limit ($\beta \to 0$) we recover the completely random transition (9.18), and in the low-temperature limit ($\beta \to \infty$) we recover the maximum information transition (9.19). In order to compute the stationary distribution, we consider the stochastic matrix describing the global dynamics (the rows denote (x_{t-1}, y_{t-1}), the columns (x_t, y_t) and the entries the transition probabilities from a state at time $t-1$ to time t):

	$(-1,-1)$	$(+1,-1)$	$(-1,+1)$	$(+1,+1)$
$(-1,-1)$	a^2	ab	ab	b^2
$(+1,-1)$	a^2	ab	ab	b^2
$(-1,+1)$	b^2	ab	ab	a^2
$(+1,+1)$	b^2	ab	ab	a^2

where

$$a := \frac{1}{1+e^{2\beta}}, \qquad b := \frac{1}{1+e^{-2\beta}}.$$

The stationary distribution is given by

$$p(+1,+1) = p(-1,-1) = \frac{1}{2} - ab, \qquad p(-1,+1) = p(+1,-1) = ab.$$

Obviously, for $\beta = 0$ we have uniform distribution $p(x,y) = \frac{1}{4}$, which implies that there is no correlation between the two nodes. As β increases, we get more and more redundancy, and in the limit $\beta \to \infty$ we get totally correlated nodes with distribution $\frac{1}{2}\left(\delta_{(-1,-1)} + \delta_{(+1,+1)}\right)$. This redundancy increase allows for compensating information about y_{t-1} by x_{t-1} and computing x_k on the basis of this information. More precisely, the two distributions $p(x_t|x_{t-1}, y_{t-1})$ and $p(x_t|x_t)$ come closer to each other. Therefore, the conditional mutual information $I(X_t; Y_{t-1} \,|\, X_{t-1})$, which is an upper bound of the transfer entropy $T(Y^{t-1} \to X_t)$, converges to zero. This fact shows that although there is no arrow from X_{t-1} to X_t, the conditional distribution $p(x_t|x_{t-1})$ is effectively dependent on x_{t-1}, which appears causally inconsistent. In order to derive a corresponding causal variant, we consider the conditional mutual information:

$$\begin{aligned}
&I(X_t; Y_{t-1} \,|\, X_{t-1}) \\
&= \sum_{x_{t-1}} p(x_{t-1}) \sum_{y_{t-1}} p(y_{t-1} \,|\, x_{t-1}) \sum_{x_t} p(x_t \,|\, x_{t-1}, y_{t-1}) \ln \frac{p(x_t \,|\, x_{t-1}, y_{t-1})}{\sum_{y'_{t-1}} p(y'_{t-1} \,|\, x_{t-1})\, p(x_t \,|\, x_{t-1}, y'_{t-1})} \\
&= \sum_{x_{t-1}} p(x_{t-1}) \sum_{y_{t-1}} p(y_{t-1} \,|\, x_{t-1}) \sum_{x_t} p(x_t \,|\, y_{t-1}) \ln \frac{p(x_t \,|\, y_{t-1})}{\sum_{y'_{t-1}} p(y'_{t-1} \,|\, x_{t-1})\, p(x_t \,|\, y'_{t-1})} \qquad (9.20)
\end{aligned}$$

With an abuse of terminology we refer to this conditional mutual information also as *transfer entropy*, which is plotted in Figure 9.12. Replacing all conditional probabilities in (9.20) by corresponding interventional ones leads to $I(X_t; Y_{t-1})$ as a causal variant of the above measure $I(X_t; Y_{t-1} \,|\, X_{t-1})$, which we refer to as *information flow*. More precisely, we have

$$\sum_{x_{t-1}} p(x_{t-1}) \sum_{y_{t-1}} p(y_{t-1} \,|\, do(x_{t-1})) \sum_{x_t} p(x_t \,|\, do(y_{t-1})) \times$$

$$\ln \frac{p(x_t \,|\, do(y_{t-1}))}{\sum_{y'_{t-1}} p(y'_{t-1} \,|\, do(x_{t-1})) p(x_t \,|\, do(y_{t-1}))}$$

$$= \sum_{x_{t-1}} p(x_{t-1}) \sum_{y_{t-1}} p(y_{t-1}) \sum_{x_t} p(x_t \,|\, y_{t-1}) \ln \frac{p(x_t \,|\, y_{t-1})}{\sum_{y'_{t-1}} p(y'_{t-1}) \, p(x_t \,|\, y'_{t-1})}$$

$$= \sum_{y_{t-1}} p(y_{t-1}) \sum_{x_t} p(x_t \,|\, y_{t-1}) \ln \frac{p(x_t \,|\, y_{t-1})}{\sum_{y'_{t-1}} p(y'_{t-1}) \, p(x_t \,|\, y'_{t-1})}$$

$$= I(X_t; Y_{t-1})$$

Comparing these two quantities, we have

$$I(X_t; Y_{t-1}) - I(X_t; Y_{t-1} \,|\, X_{t-1}) \;=\; I(X_t; X_{t-1}) \;\geq\; 0.$$

Intuitively speaking, in this example, the transfer entropy captures only one part of the causal information flow.

The following diagram shows the shape of the conditional mutual information and the information flow as function of β. As we see, the information flow is consistent with the intuition that moving from $\beta = 0$ to $\beta = \infty$ corresponds to an interpolation between a transition with vanishing information flow and a transition with maximal information flow. Near $\beta = 0$ the transfer entropy increases as β becomes larger and is close to the information flow. But for larger β's it starts decreasing and converges to zero for $\beta \to \infty$. The reason for that is simply that the transition for large β generates more redundancy between the two processes X and Y. Therefore, as β grows, an increasing amount of information about Y_{t-1} can be computed from information about X_{t-1}, which lets the transfer entropy decrease towards zero.

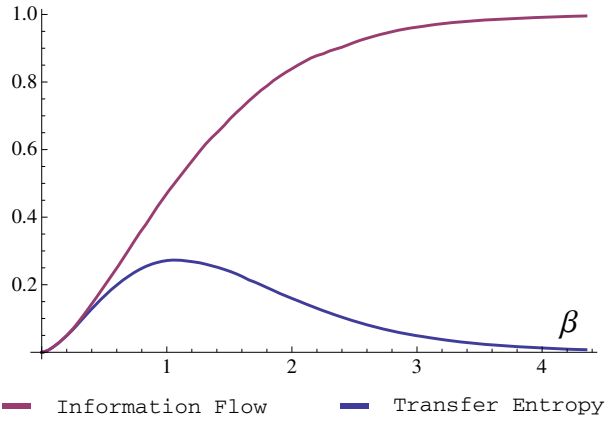


Fig. 9.12 Transfer entropy and information flow

9.4.3 Information Flows in the Sensorimotor Loop

In this section, we apply the notion of transfer entropy to the context of the sensorimotor loop. Sporns and Lungarella used the transfer entropy in order to describe information flows through the sensorimotor loop (Lungarella and Sporns 2006). Here, we approach this subject from a theoretical perspective and point out that one has to be very careful with the interpretation of transfer entropy as a causal measure. Our considerations will be based on the causal diagram shown in Figure 9.13. Here, we have four processes W_t, S_t, C_t, A_t, $t \geq 1$, and the initial node H which stands for "history."

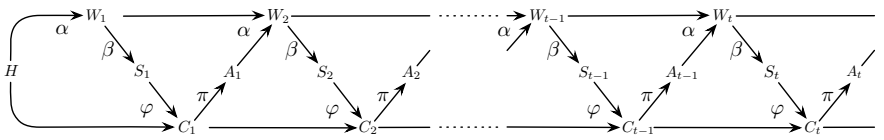


Fig. 9.13 Causal diagram of the sensorimotor loop

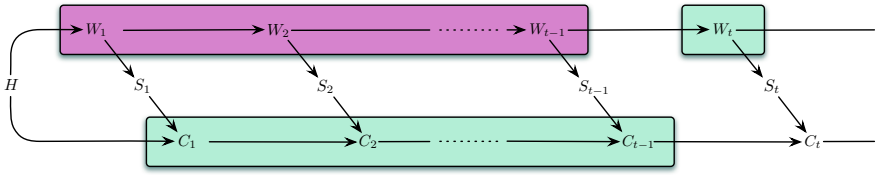
Let us study the two simple cases of a passive observer and an open loop controller (see Figures 9.5 and 9.6). In these cases, some of the arrows are missing, and therefore we know from the definition of the respective causal structure that particular causal effects are not present. We can then evaluate the corresponding transfer entropy and thereby test its consistency with the given causal structure. We start with the passive observer, shown in Figure 9.5. Here, the actuator with its arrows is removed, which means that the agent is not able to act on the world. Therefore, we expect $T(C^{t-1} \to W_t) = 0$ and $T(S^{t-1} \to W_t) = 0$. This is confirmed by the fact that C^{t-1} and W_t are d-separated by W^{t-1} (see Figure 9.14). Similarly, S^{t-1} and W_t are d-separated by W^{t-1}.

Now let us consider the open loop controller. Here, the sensor with its arrows is removed, and we expect no causal effect of W on C and also no causal effect of A on C. Therefore, consistency of the transfer entropy with the causal structure would require $T(W^{t-1} \to C_t) = 0$ and $T(A^{t-1} \to C_t) = 0$. This is confirmed in terms of the same d-separation arguments as in the above case of a passive observer (see Figure 9.15).

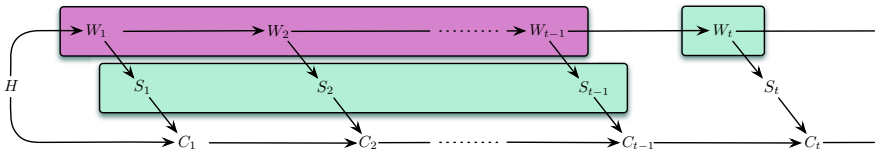
Summarising the above considerations, we have the following proposition.

Proposition 3. *For the passive observer one has* $T(C^{t-1} \to W_t) = 0$ *and* $T(S^{t-1} \to W_t) = 0$ *for all* $t \geq 1$. *In the context of open loop control we have* $T(W^{t-1} \to C_t) = 0$ *and* $T(A^{t-1} \to C_t) = 0$ *for all* $t \geq 1$.

Proposition 3 confirms our expectation that a passive observer has no causal effect on the world, and that within open loop control there is no causal effect of the world on the agent (note that, with "agent" we mean the internal process C of the agent). However, one has to be careful with this measure. For example, in the situation of a passive observer, there is no directed path from C to S. Therefore, one

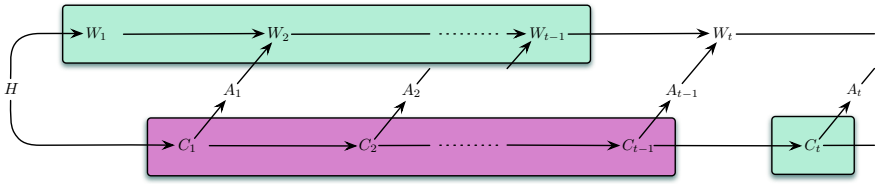


(A)

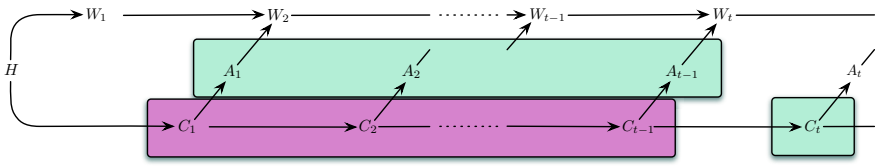


(B)

Fig. 9.14 Sensorimotor loop of a passive observer. The diagram (A) shows that $T(C^{t-1} \to W_t) = 0$, and diagram (B) shows $T(S^{t-1} \to W_t) = 0$.



(A)



(B)

Fig. 9.15 Sensorimotor loop of an open loop controller. The diagram (A) shows that $T(W^{t-1} \to C_t) = 0$, and diagram (B) shows $T(A^{t-1} \to C_t) = 0$.

would not only expect the absence of causal effects of the agent on the world but also the absence of causal effects on the sensor process, which is obtained from the world process. In other words, it is natural to have not only $T(C^{t-1} \to W_t) = 0$, which is confirmed in Proposition 3, but also $T(C^{t-1} \to S_t) = 0$. From the causal structure

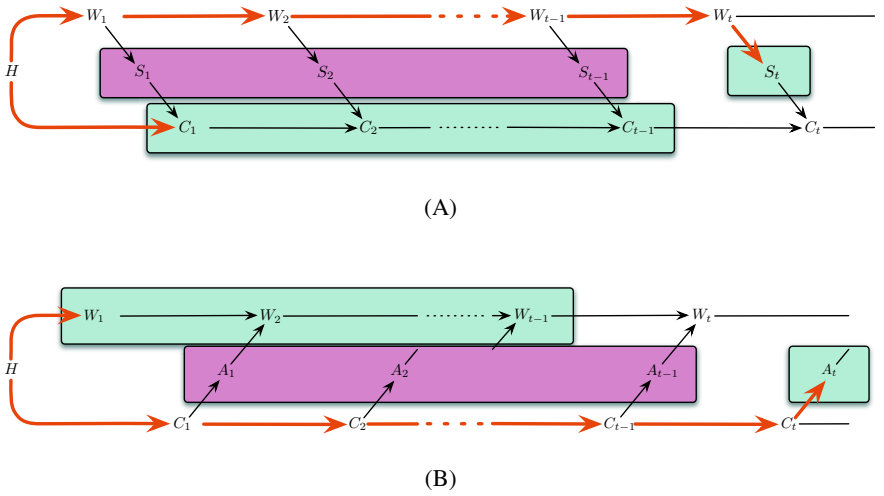


Fig. 9.16 (A) Sensorimotor loop of a passive observer. Although C^{t-1} is not a cause of S_t, that is $C^{t-1} \not\to S_t$, it is indeed possible that $T(C^{t-1} \to S_t) > 0$. This is seen by the fact that the sets C^{t-1} and S_t are not d-separated by S^{t-1} (see unblocked path). (B) Sensorimotor loop of an open loop controller. Although W^{t-1} is not a cause of A_t, that is $W^{t-1} \not\to A_t$, it is indeed possible that $T(W^{t-1} \to A_t) > 0$. This is seen by the fact that the sets W^{t-1} and A_t are not d-separated by A^{t-1} (see unblocked path).

shown in Figure 9.16 (A) this is not confirmed in terms of the d-separation criterion. More precisely, S_t and C^{t-1} are not d-separated by S^{t-1}, and $T(C^{t-1} \to S_t) > 0$ is indeed possible, which appears counterintuitive. We have the same problem with open loop control (see Figure 9.16 (B)). In this case, one would expect that there is no causal effect of the world on the agent's actions, that is $T(W^{t-1} \to A_t) = 0$. However, A_t and W_{t-1} are not d-separated by A^{t-1}, and, again, $T(W^{t-1} \to A_t) > 0$ is possible although W^{t-1} is not a cause of A_t.

A further method to study information flows, which has been proposed in (Ay and Polani 2008), is based on the interventional calculus of conditioning as already applied at the end of Section 9.4.2. Here, we briefly outline how this method can be applied within the context of the sensorimotor loop. In order to do so, we consider the causal effects listed in Proposition 2. The information flow from C to S, for instance, can be quantified by

$$IF(C \to S) := \sum_c p(c) \sum_s p(s|do(c)) \ln \frac{p(s|do(c))}{\sum_{c'} p(c')\, p(s|do(c'))} \tag{9.21}$$

which is a causal variant of the mutual information

$$I(C;S) = \sum_c p(c) \sum_s p(s|c) \ln \frac{p(s|c)}{\sum_{c'} p(c')\, p(s|c')}.$$

This expression (9.21) requires an intervention in C, which will, in general, disturb the system. Therefore, it is not clear to what extent this quantity refers to the in situ situation of a system while functioning. It is possible that information flow patterns change as consequence of intervention. One attempt to resolve this problem is given by the method of virtual intervention. Here, one predicts the consequences of interventions based on observations only, without actually applying an interventional operation. The system remains unperturbed, and one can still evaluate the expression (9.21). To be more precise, we use Proposition 2 and replace in (9.21) all distributions $p(s|do(c))$ by $\sum_a p(a|c)\sum_{c'} p(s|c',a)\,p(c')$. This leads to an expression of the information flow $IF(C \to S)$ that does not involve any experimental intervention but only observation. In addition, all involved variables are accessible to the agent, which enables the agent to evaluate the information flow while being embedded in the sensorimotor loop. Currently, it remains unclear whether this method of virtual intervention resolves the above-mentioned problem.

9.5 Predictive Information and Its Maximization – An Experimental Case Study

In the previous Sections 9.4.2 and 9.4.3 we focussed on information flows between various interacting processes, such as the control process C on the sensor process S in the sensorimotor loop. In this section, we now concentrate on temporal flows within one single process. To this end, consider first a stochastic process X_t, $t \in \mathbb{Z}$. Furthermore, given three time points $t_- < t < t_+$, we consider the mutual information

$$I(X_{t_-},\dots,X_t;X_{t+1},\dots,X_{t_+}) \;=\; H(X_{t+1},\dots,X_{t_+}) - H(X_{t+1},\dots,X_{t_+}\,|\,X_{t_-},\dots,X_t).$$

If we assume that the variables $X_{t_-},\dots,X_t$ represent the past (and present) of observed variables and $X_{t+1},\dots,X_{t_+}$ the future or unobserved variables with respect to t, then this mutual information quantifies the reduction of uncertainty about the future given the past. In other words, it quantifies the amount of information in the future that can be predicted in terms of past observations. Therefore, the corresponding limit for $t_- \uparrow \infty$ and $t_+ \uparrow \infty$, which always exists but can be infinite, is referred to as *predictive information* (PI) (Bialek et al. 2001). It is also known as *excess entropy* (Crutchfield and Young 1989) and *effective measure complexity* (Grassberger 1986). Note that the predictive information is independent of t if the process is stationary.

In the context of the sensorimotor loop, the predictive information and related quantities serve as objective functions for self-organized learning. Of particular interest is the predictive information of the sensor process $S = (S_t)_{t\in\mathbb{Z}}$, which quantifies the amount of information in the agent's future sensor process that can be predicted by the agent based on the observation of its past (and present) sensor process. For simplicity, we also consider the lower bound $I(S_t;S_{t+1})$ of the predictive information and, with abuse of terminology, we refer to this simplified quantity also

as predictive information (PI). Clearly, the predictive information is, on one hand, dependent on the policy and, on the other hand, also dependent on the embodiment of the agent. Policies with high predictive information correspond to niches of predictability within the sensorimotor loop and allow the agent to exploit these niches for task oriented behavior. Learning processes that maximize several variants of predictive information have been proposed and studied in view of their behavioral implications (Ay et al. 2008; Zahedi et al. 2010; Ay et al. 2012; Martius et al. 2013).

Here, we focus on the experimental case study of our previous work (Zahedi et al. 2010). In our experiments, embodied agents maximize the predictive information calculated on their sensor data by modulation of their policies. Different controller types are evaluated and the results are coordinated behaviors of passively coupled, individually controlled robots. From these results, three conclusions will be drawn and discussed. First, the different controller structures lead to a conclusion about optimal design, second, PI maximization leads to the formation of behavioral modes, and third, PI maximization leads to morphological computation (Pfeifer and Bongard 2006).

For the implementation of the learning rule in an embodied system, we chose a discrete-valued representation of the probability distributions for the following reason. At this initial step of evaluating the PI as a self-organised learning principle, we wanted to use as few assumptions as possible about the underlying model. Implementing the learning rule in the continuous domain generally requires more assumptions and restrictions, as the following example demonstrates. An implementation of the policy as a (recurrent) neural network binds the space of possible functions to the structure of the neural network (Pasemann 2002). Changing the weights and biases of the network only permits variation among the functions defined by the structure (and neural models). To avoid such a pruning in the space of policies, we chose stochastic matrices. To reduce the complexity, we first concentrated on reactive control (see Figure 9.7). Given a world model γ and a policy π, an estimate of the PI can be calculated with intrinsically available information by the following set of equations

$$I^{(\gamma)}(S_t;S_{t+1}) = \sum_{s_t,s_{t+1}\in\mathcal{S}} p(s_t,s_{t+1}) \ln \frac{p(s_t,s_{t+1})}{p(s_t)\,p(s_{t+1})} \tag{9.22}$$

$$p(s_t,s_{t+1}) = \sum_{a_t\in\mathcal{A}} p(s_t,a_t,s_{t+1}) = \sum_{a_t\in\mathcal{A}} p(s_t)\,\pi(s_t;a_t)\,\gamma(s_t,a_t;s_{t+1}) \tag{9.23}$$

$$p(s_t) = \sum_{s_{t+1}\in\mathcal{S}} p(s_t,s_{t+1}), \qquad p(s_{t+1}) = \sum_{s_t\in\mathcal{S}} p(s_t,s_{t+1}) \tag{9.24}$$

which can be used to calculate a *natural gradient* iteration (Amari 1998) of the PI with respect to the policy $\pi(s_t;a_t)$ in the following way (Zahedi et al. 2010)

$$\pi^{(0)}(s;a) := \frac{1}{|\mathscr{S}|} \qquad n \in \mathbb{N}\backslash\{0\}$$

$$\pi^{(n)}(s;a) = \pi^{(n-1)}(s;a) + \frac{1}{n+1}\,\pi^{(n)}(s;a)\left(F(s) - \sum_a \pi^{(n-1)}(s;a)F(s)\right) \tag{9.25}$$

$$F(s) := p^{(n)}(s)\sum_{s'} \gamma^{(n)}(s,a;s')\log_2 \frac{\sum_a \pi^{(n-1)}(s;a)\gamma^{(n)}(s,a;s')}{\sum_{s''} p^{(n)}(s'')\sum_a \pi^{(n-1)}(s'';a)\,\gamma^{(n)}(s'',a;s')}$$

The sensor distribution $p(s_t)$ and the intrinsic world model $\gamma(s_t,a_t;s_{t+1})$ are sampled according to

$$p^{(0)}(s) := \frac{1}{|\mathscr{S}|}$$

$$p^{(n)}(s) := \begin{cases} \dfrac{n}{n+1}\,p^{(n-1)}(s) + \dfrac{1}{n+1} & \text{if } S_{n+1} = s \\ \dfrac{n}{n+1}\,p^{(n-1)}(s) & \text{if } S_{n+1} \neq s \end{cases} \tag{9.26}$$

$$\gamma^{(0)}(s,a;s') := \frac{1}{|\mathscr{S}|}$$

$$\gamma^{(n_a^s)}(s,a;s') := \begin{cases} \dfrac{n_a^s}{n_a^s+1}\,\gamma^{(n_a^s-1)}(s,a;s') + \dfrac{1}{n_a^s+1} & \text{if } S_{n_a^s+1} = s',\ S_n = s,\ A_{n_a^s+1} = a \\ \dfrac{n_a^s}{n_a^s+1}\,\gamma^{(n_a^s-1)}(s,a;s') & \text{if } S_{n_a^s+1} \neq s',\ S_n = s,\ A_{n_a^s+1} = a \\ \gamma^{(n_a^s-1)}(s,a;s') & \text{if } S_{n_a^s} \neq s \text{ or } A_{n_{s,a}+1} \neq a \end{cases} \tag{9.27}$$

This concludes the brief presentation of the learning rule. See (Zahedi et al. 2010) for a detailed description. The next paragraphs describe the experiments and discuss the results.

All experiments were conducted purely in simulation for the sake of simplicity, speed and analysis. Current simulators, such as YARS (Zahedi et al. 2008), which was chosen for the experiments presented below, are shown to be sufficiently realistic to simulate the relevant physical properties of mobile robots, and designed such that experimental runs can be automated, run at faster than real-time speed, and require minimum effort to set-up.

The experiments are conducted with differential wheeled robots (see Figure 9.17A). The only sensors of each robot are the current wheel velocities S_t and the only actions are the motor commands for the next wheel velocities A_t. Each robot is either controlled by a single controller (see Figure 9.17C) or by two controllers (see Figure 9.17D). We refer to the single controller setting as *combined* control and to the two controller setting as *split* control.

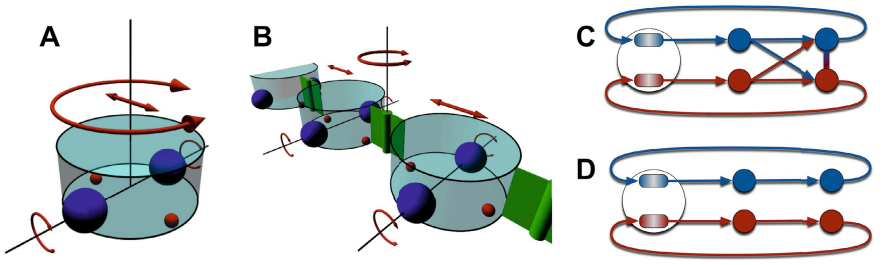


Fig. 9.17 Experimental set-up. A) Single circular differential wheeled robot. The image shows the two driving wheels and the possible movements that they allow for. B) This image shows the passive coupling between the robots. C) This image is a schema of the combined control, where one controller reads the information of both wheels and controls the velocity of both wheels. D) In the split controller configuration, each wheel has a controller that is independent of the other. Hence, each controller only reads the information of one wheel and controls the velocity of only one wheel.

Three and five robots are passively connected to a chain of robots (see Figure 9.17B), which results in the following four experiments:

1. three robots with combined control,
2. three robots with split control,
3. five robots with combined control, and
4. five robots with split control.

We will refer to these settings by the number of robots and the controller type. This means that 3C, 3S, 5C, and 5S refer to three robots with combined control, three robots with split control, etc.

The results of the experiments (see Figure 9.18) show that all systems maximize the PI and that they all perform some sort of exploration behavior. As the maximization of the PI does not specify a behavior of a system that can be evaluated directly, we need to define a measure based on our observations. In a chain of individually controlled robots, the travelled distance over time is a good indication for the quality of the coordination, as only well coordinated robots will be able to travel far. This is why we chose a sliding window coverage entropy (see Figure 9.18 and (Zahedi et al. 2010) for details) to measure the exploration behavior. It must be noted, that we are not interested in the exploration itself, but rather in the quantifiable, qualitative change of the observable behavior of the systems which result from the PI maximization. In this context, two counter-intuitive results are shown in Figure 9.18. First, the robot chains with five robots outperform the robot chains with three robots (compare 5C with 3C and 5S with 3S in Figure 9.18), and second, the split control systems outperform the corresponding combined control systems (compare 5S with 5C and 3S with 3C in Figure 9.18), in both, maximizing the PI and maximizing the coordination of the robots in a chain.

The first result is counter-intuitive, because longer chains means that more robots, and hence, more controllers have to coordinate based on the local information

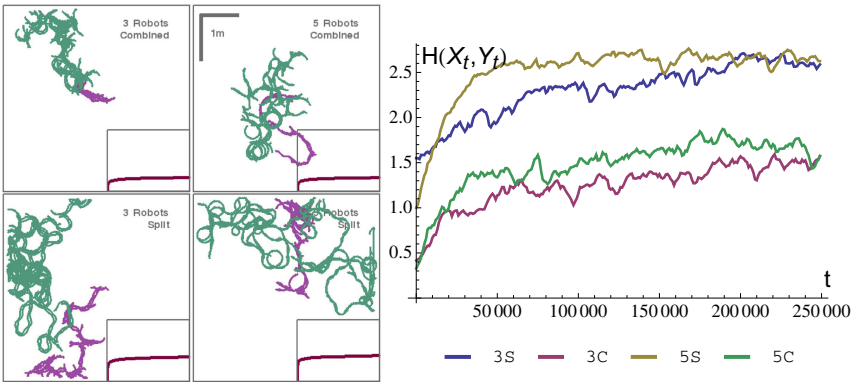


Fig. 9.18 PI maximization results. The graph on the left-hand shows four plots with one sub-plot each. The trajectory plots show the initial ten minutes (purple) and the final ten minutes (green) of the behavior (after 106 policy updates, which is approximately 27 hours). The sub-plots show the progress of the PI averaged over 100 runs for the same amount of time, normalised to the unit interval. The plot on the right-hand side shows how the exploration quality of the controllers progresses over time. For this purpose, the entropy $H(X_t, Y_t)$ over the spatial coordinates X_t, Y_t is calculated for a sliding window (see (Zahedi et al. 2010) for a discussion). Both plots show, indirectly measured by the sliding window coverage entropy, that the chains with five robots show a higher coordination compared to those with three robots, and that split control results in a higher coordination compared to combined control.

available through the wheel velocity sensors only. The second result is counter-intuitive, because the combined controller should, if anything, have additional properties compared to the split control as it combines the sensor information of both wheels. The next paragraph will analyse the behavior of one representative policy of the 5S setting and thereby explain all the results.

The behavior of all robots can be categorised into three modes. The first mode is called *forward* movement and it is characterised by (mainly) positive wheel velocities. The second mode is called *backward* movement and it is analogously characterised by (mainly) negative wheel velocities. The third is a transition between the two previous modes in which there is no clear direction of the robot chain.

To understand how the modes occur, we recorded the data stream of the sensors and actions ($s(t)$ and $a(t)$) of one representative controller of the 5S setting. According to the update rules used for the world model (see Eq. 9.27) we sampled the following four conditional probabilities $p(A_t \geq 0|S_t \geq 0)$, $p(A_t < 0|S_t \geq 0)$, $p(A_t < 0|S_t < 0)$, and $p(A_t \geq 0|S_t < 0)$, where the first two and the latter two sum up to one (see Figure 9.19). The two conditional probabilities $p(A_t \geq 0|S_t \geq 0)$ and $p(A_t < 0|S_t < 0)$ refer to maintaining the current direction of travel, whereas, $p(A_t < 0|S_t \geq 0)$ and $p(A_t \geq 0|S_t < 0)$ refer to a switching of the current direction of travel. At the beginning of the learning process, all conditional probabilities are by definition equal (see Eq. 9.25). After a while, it is seen that the system is more likely to maintain its current movement compared to switching as the sign of the

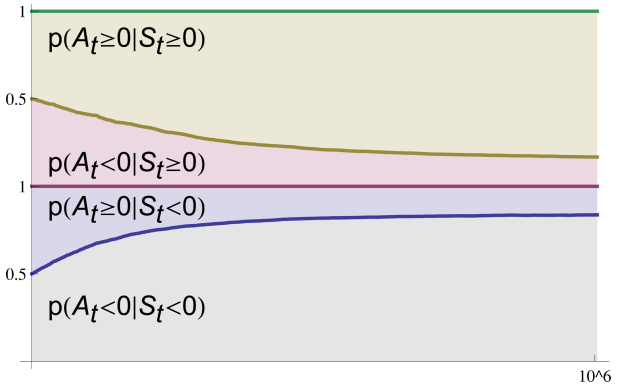


Fig. 9.19 Formation of modes. This graph shows the four conditional probabilities $p(A_t \geq 0|S_t \geq 0)$, $p(A_t < 0|S_t \geq 0)$, $p(A_t < 0|S_t < 0)$, and $p(A_t \geq 0|S_t < 0)$ plotted above each other over time and summed accordingly. The plot reveals that the policies are learned such that they maintain their current mode (forward or backward) with a high probability and that switching between the two modes occurs with the same probability. This fulfils the requirements of diversity and compliance posed by the PI.

action $a(t)$ is more likely chosen to be equal to the sign of the sensed wheel velocity $s(t)$ (see Figure 9.19). At the end of the learning process (10^6 iterations), the conditional probabilities are approximately $p(A_t \geq 0|S_t \geq 0) \approx p(A_t < 0|S_t < 0) \approx 0.85$ and $p(A_t < 0|S_t \geq 0) \approx p(A_t \geq 0|S_t < 0) \approx 0.15$. From these estimations, we can now reconstruct how the modes occur. If a robot chain is currently in the forward mode ($S_t \geq 0$ for all wheels), it requires more than half the controllers to decide on switching the mode for the robot chain to change its direction of movement. For the chain with three robots, it requires at least four controllers, and for the chain with five robots, six controllers to decide to switch directions. Hence, the probability of switching, denoted by p_s, is $p_s \leq 0.15^4$ for the three robot chain and $p_s \leq 0.15^6$ for the five robot configuration, where the probability refers to the controller update frequency (10Hz). For the two chains of robots, this means that the overall probability of maintaining the current behavioral mode in every time step is larger than $1 - 0.15^4 = 99.949\%$ for 3S and larger than $1 - 0.15^6 = 99.999\%$ for 5S. This explains why longer chains outperform shorter chains in terms of exploration as they are more likely to maintain their current direction of movement. The modes of the 5S are more distinctive compare to the 3S due to the larger number of robots in the chain. This can also be considered as a form of morphological computation (Pfeifer and Bongard 2006), which we will address later in this chapter again. The next question to answer is why the modes are beneficial in term of maximizing the PI. From the discussion above it follows that $H(S_{t+1}|S_t)$ is minimized because knowledge of the current wheel velocity reduces the uncertainty of the next wheel velocity significantly due to the formation of the modes. As the switching probabilities are almost equal, all sensor states are equally often perceived, which maximizes the entropy

$H(S_{t+1})$. This means the system shows a compliant variance in its behavior as it is demanded by the PI.

The second counter-intuitive result was that split controllers outperform combined controllers in exploration and PI maximization (see Figure 9.18). This is counter-intuitive because the combined controller has additional features (compare Figure 9.17C with Figure 9.17D)) compared to two split controllers. If the split controllers are likely to find the good or optimal solutions, then the combined controllers should be able to

1. find the same good or optimal solutions,
2. find other good or optimal solutions, and
3. find even better solutions.

The question is, why is this not the case? The space of possible policies spanned by the two split controllers is a subspace of possible controllers spanned by one combined controller. It happens to be that this subset of the split controllers encloses only a few maximizers and that a sufficiently large number of them is optimal with respect to the maximization of the PI. This is more obvious if only one robot is allowed to learn with two split and one combined controller. This is not shown here but discussed in detail in (Zahedi et al. 2010). Due to the low number of parameters defining the subspace of the split controllers, the optimisers are found faster and more reliably. In the superset of the combined controller, we find many suboptimal solutions which are more likely to be found compared to the optimisers. This means, by splitting the controllers we have made a large subspace of the combined controller space inaccessible to learning. The resulting subspace still had all the maximizers of the PI which is why the split controller outperforms the combined controller. Concluding, if one finds a natural way to restrict the policy space (possibly according to the morphology), such that it captures all maximizers of a given function, then this would be called an *optimal control* as the policy is optimally parameterised for learning and control. Finding such a natural method is an ongoing topic in this field of research.

As stated earlier, we want to discuss the results also from the perspective of morphological computation (Pfeifer and Bongard 2006). We already saw that the number of robots in the chain directly influences the exploration behavior. That the maximization of the PI leads to morphological computation is more obvious, if we take the 5S system, and remove the passive joints between the robots (see Figure 9.20). Both plots in the figure show the trajectories of all 10 wheels of the five robots, where the wheels of each robot share one color. The left-hand side shows the exploration behavior that we have already seen (compare with Figure 9.18). The right-hand side of the figure shows that the uncoupled robots loose a lot of their original behavior. They rotate more often on the spot and the trajectories are not as long and smooth as in the coupled system.

As the controllers are identical in both settings, this means that there is a contribution of the world (morphology and environment) to the behavior which cannot be assigned to the controller as they are identical in both systems. This contribution is called *morphological computation* (Pfeifer and Bongard 2006). Various

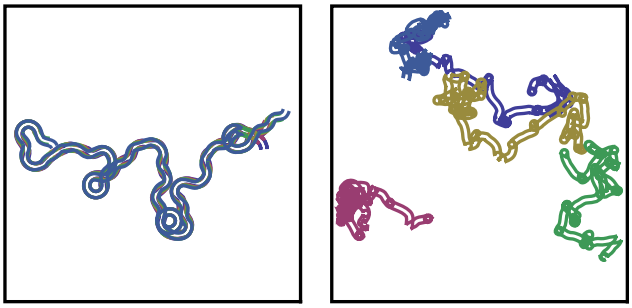


Fig. 9.20 Morphological computation. The two plots show the trajectories of all wheels of all five robots of the split control setting for ten seconds. The wheels of one robot share the same color in the plots. The plot on the left-hand side shows the trajectories of the original, passively coupled robots. The plot on the right-hand side shows the same five robots with the same ten controllers but with the passive connections removed. The comparison of the plots shows that the behavior changes significantly, which leads to two conclusions. First, PI maximization adapts to the world, and second, PI maximization leads to morphological computation, as the behavior is also significantly determined by the morphology of the system.

quantifications of morphological computation are derived and evaluated in experiments in (Zahedi and Ay 2013). They are based on causal and associative information-theoretic measures.

Appendix

Proof of Proposition 2:
(1)

$$p(h,c,w,s\,|\,do(a)) = p(h)\,\varphi(h;c)\,\alpha(h,a;w)\,\beta(w;s).$$

This implies

$$p(s,c\,|\,do(a)) = \sum_{h,w} p(h)\,\varphi(h;c)\,\alpha(h,a;w)\,\beta(w;s)$$

$$\begin{aligned} p(c\,|\,do(a)) &= \sum_{s}\sum_{h,w} p(h)\,\varphi(h;c)\,\alpha(h,a;w)\,\beta(w;s) \\ &= p(c) \end{aligned}$$

$$
\begin{aligned}
p(s\,|\,do(a),c) &= \frac{p(s,c\,|\,do(a))}{p(c\,|\,do(a))} \\
&= \sum_{h,w} \tfrac{p(h)}{p(c)}\, \varphi(h;c)\, \alpha(h,a;w)\, \beta(w;s) \\
&= \sum_{h,w} p(h\,|\,c)\, p(w\,|\,h,a)\, p(s\,|\,w) \\
&= \sum_{h,w} p(h\,|\,c,a)\, p(w\,|\,h,a,c)\, p(s\,|\,w,h,a,c) \\
&\quad \text{(conditional independence, see diagram in Figure 9.11)} \\
&= p(s\,|\,a,c).
\end{aligned}
$$

The second and third equations of the proposition follow from the general theory (see (Pearl 2000), Theorem 3.2.2 (Adjustment for Direct Causes), and Theorem 3.3.4 (Front-Door Adjustment)). For completeness, we prove them directly.

(2)

$$
\begin{aligned}
p(s\,|\,do(a)) &= \sum_{h,c,w} p(h,c,w,s\,|\,do(a)) \\
&= \sum_{h,c,w} p(h)\, \varphi(h;c)\, \pi(c;a)\, \alpha(h,a;w)\, \beta(w;s)\, \frac{1}{p(a|c)} \\
&= \sum_{h,c,w} \frac{p(h,c,a,w,s)}{p(c,a)}\, p(c) \\
&= \sum_{c} p(s|c,a)\, p(c).
\end{aligned}
$$

(3)

$$
\begin{aligned}
p(s\,|\,do(c)) &= \sum_{h,a,w} p(h,a,w,s\,|\,do(c)) \\
&= \sum_{a} \pi(c;a) \sum_{h,w} p(h)\, \alpha(h,a;w)\, \beta(w;s) \\
&= \sum_{a} p(a|c) \sum_{h,w} \left(\sum_{c'} p(c')\, p(h|c') \right) p(w|h,a)\, p(s|w) \\
&= \sum_{a} p(a|c) \sum_{c'} p(c') \sum_{h,w} p(h|c')\, p(w|h,a)\, p(s|w) \\
&= \sum_{a} p(a|c) \sum_{c'} p(c') \sum_{h,w} p(h|c',a)\, p(w|h,a,c')\, p(s|w) \\
&= \sum_{a} p(a|c) \sum_{c'} p(c')\, p(s|c',a).
\end{aligned}
$$

□

Acknowledgements. This work has been supported by the DFG Priority Program 1527, *Autonomous Learning*, and by the Santa Fe Institute.

References

Amari, S.: Natural gradient works efficiently in learning. Neural Computation 10(2), 251–276 (1998)

Ay, N., Bernigau, H., Der, R., Prokopenko, M.: Information driven self-organization: The dynamical system approach to autonomous robot behavior. Theory Biosci. (131), 161–179 (2012)

Ay, N., Bertschinger, N., Der, R., Güttler, F., Olbrich, E.: Predictive information and explorative behavior of autonomous robots. EPJ B 63(3), 329–339 (2008)

Ay, N., Polani, D.: Information flows in causal networks. Advances in Complex Systems 11(1), 17–41 (2008)

Bialek, W., Nemenman, I., Tishby, N.: Predictability, complexity, and learning. Neural Computation 13(11), 2409–2463 (2001)

Crutchfield, J.P., Young, K.: Inferring statistical complexity. Phys. Rev. Lett. 63(2), 105–108 (1989)

Der, R., Güttler, F., Ay, N.: Predictive information and emergent cooperativity in a chain of mobile robots. In: ALife XI. MIT Press (2008)

Grassberger, P.: Toward a quantitative theory of self-generated complexity. International Journal of Theoretical Physics 25(9), 907–938 (1986)

Kaiser, A., Schreiber, T.: Information transfer in continuous processes. Physica D: Nonlinear Phenomena 166(1-2), 43–62 (2002)

Klyubin, A.S., Polani, D., Nehaniv, C.L.: Tracking information flow through the environment: Simple cases of stigmerg. In: Pollack, J. (ed.) Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems, pp. 563–568 (2004)

Klyubin, A.S., Polani, D., Nehaniv, C.L.: Empowerment: A universal agent-centric measure of control. In: Proc. CEC. IEEE (2005)

Lauritzen, S.L.: Graphical Models. Oxford University Press (1996)

Lungarella, M., Sporns, O.: Information self-structuring: Key principle for learning and development. In: IEEE (ed.) Proc. the 4th International Conference on Development and Learning, pp. 25–30. IEEE Press, San Diego (2005)

Lungarella, M., Sporns, O.: Mapping information flow in sensorimotor networks. PLoS Comp. Biol. 2(10), e144 (2006)

Martius, G., Der, R., Ay, N.: Information driven self-organization of complex robotic behaviors. PLOS One 8(5) (2013), doi:10.1371/ journal.pone.0063400

Massey, J.L.: Causality, feedback and directed information. In: Proc. 1990 Intl. Symp. on Info. Th. and its Applications, pp. 27–30 (1990)

Pasemann, F.: Complex dynamics and the structure of small neural networks. Network: Computation in Neural Systems 13(2), 195–216 (2002)

Pearl, J.: Causality: Models, Reasoning and Inference. Cambridge University Press (2000)

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

Polani, D., Nehaniv, C., Martinetz, T., Kim, J.T.: Relevant Information in Optimized Persistence vs. Progeny Strategies. In: Rocha, L.M., Bedau, M., Floreano, D., Goldstone, R., Vespignani, A., Yaeger, L. (eds.) Proc. Artificial Life X, pp. 337–343. MIT Press, Cambridge (2006)

Prokopenko, M., Gerasimov, V., Tanev, I.: Evolving spatiotemporal coordination in a modular robotic system. In: Nolfi, S., Baldassarre, G., Calabretta, R., Hallam, J.C.T., Marocco, D., Meyer, J.-A., Miglino, O., Parisi, D. (eds.) SAB 2006. LNCS (LNAI), vol. 4095, pp. 558–569. Springer, Heidelberg (2006)

Prokopenko, M., Lizier, J.T., Price, D.C.: On thermodynamic interpretation of transfer entropy. Entropy 15(2), 524–543 (2013)

Reichenbach, H.: The Direction of Time. University of California Press (1956)

Schreiber, T.: Measuring information transfer. Physical Review Letters 85(2) (2000)

Zahedi, K., Ay, N.: Quantifying morphological computation. Entropy 15(5), 1887–1915 (2013)

Zahedi, K., Ay, N., Der, R.: Higher coordination with less control – a result of information maximization in the sensori-motor loop. Adaptive Behavior 18(3-4), 338–355 (2010)

Zahedi, K., von Twickel, A., Pasemann, F.: YARS: A physical 3D simulator for evolving controllers for real robots. In: Carpin, S., Noda, I., Pagello, E., Reggiani, M., von Stryk, O. (eds.) SIMPAR 2008. LNCS (LNAI), vol. 5325, pp. 75–86. Springer, Heidelberg (2008)