Tracking Information Flow through the Environment: Simple Cases of Stigmergy

Alexander S. Klyubin, Daniel Polani, and Chrystopher L. Nehaniv

Adaptive Systems Research Group
School of Computer Science, Faculty of Engineering and Information Sciences
University of Hertfordshire, College Lane, Hatfield Herts AL10 9AB, UK
{A.Kljubin,D.Polani,C.L.Nehaniv}@herts.ac.uk

Abstract

Recent work in sensor evolution aims at studying the perception-action loop in a formalized information-theoretic manner. By treating sensors as extracting information and actuators as having the capability to "imprint" information on the environment we can view agents as creating, maintaining and making use of various information flows. In our paper we study the perception-action loop of agents using Shannon information flows. We use information theory to track and reveal the important relationships between agents and their environment. For example, we provide an information-theoretic characterization of stigmergy and evolve finite-state automata as agent controllers to engage in stigmergic communication. Our analysis of the evolved automata and the information flow provides insight into how evolution organizes sensoric information acquisition, implicit internal and external memory, processing and action selection.

1 Introduction

We approach the study of information flows through the perception-action loop and the environment using classical information theory. We believe that a formalized approach to the perception-action loop may bring us closer to finding principles underlying adaptive behavior. Such principles could be used both for guiding the construction of systems with desired information flows and for studying their behavior. The use of information theory provides us with a universal framework which minimizes the influences of a particular implementation.

Consider an agent with sensors and actuators. Sensors capture some information, the information gets processed, and based on the results the actuators act upon the environment. If sensors are seen as taking information in, it seems also reasonable to see actuators as modifying the environment informationally. Surprisingly, it seems that little research has been done to quantitatively treat perception-action in terms of information – an observation also made by Touchette and Lloyd in the context of control [15].

In [14, 15] the problem of control is quantitatively treated in terms of Shannon information which is seen as flowing from a controlled system into a controller and then back. An important information-theoretic bound is obtained for the usefulness of any sensor for control. [9] introduces an information-theoretic view of perception and actuation and discusses usefulness as a means to attribute agent-specific meaning to information. This is further formalized in the context of relevant information [12] measured in bits. Relevant information "flows" from the environment via sensors to actuators, thus connecting them.

1.1 Shannon Information Flow & Environment

One of the motivations for our study is sensory evolution [5]. Originally the idea was focused on sensors evolving to capture more useful or relevant information. However, using sensors is often inseparable from what an agent needs to *do* in its environment. It makes more sense to consider perception and action together as a single entity, a loop. Actions in the environment can influence the sensors, creating a loop. This is an important link which the quantitative approaches above have not addressed directly and which this paper does address in a quantitative manner (Fig. 1).

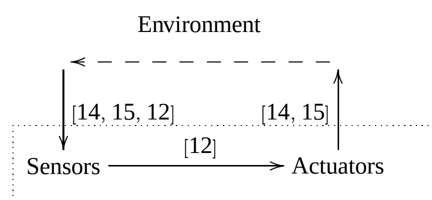

Figure 1: Information flow within an agent (box) and between the agent and its environment addressed by quantitative information-theoretic approaches. Solid line – prior work. Dashed and solid – this paper.

The perception-action *loop* is important for understanding the behavior of adapted agents. If we view sensors as capturing Shannon information and actuators as being capable of "imprinting" information on the environment, we can then treat agents as creating, maintaining and using various information flows, both internal and external. The view may be quite useful since there are strong indications that biological agents are partly driven by the necessity of acquiring, exchanging and also concealing information.

Treating the perception-action loop in terms of Shannon information flows enables us to quantitatively capture certain phenomena, like "imprinting" information onto the environment, offloading and later reacquisition of information, and active perception. An agent imprints information onto its environment by changing the environment in different ways depending on its internal state.

Uncovering hidden information using the perception-action loop is demonstrated by Kirsh and Maglio in [8]. There advanced Tetris game players are shown to quickly rotate a falling block while it still is not visible completely. This active modification of the environment allows the players to discover the shape of the block before it actually becomes completely visible on the screen. *Active perception* in general is seen as a powerful technique [10].

Offloading of information into the environment can be illustrated by the fact that we often write notes or reminders, which we then later look at to reacquire some information. A good account of such information flow (not in the Shannon sense) between several people is given in the analysis of how members of an airliner crew indirectly communicate using cockpit controls as a medium [7]. For example, long before landing one of the pilots calculates proper flap settings for various speeds and then based on the results sets special markers on the airspeed indicator. Later, during the landing phase, the markers allow the crew to quickly and reliably find out what flap settings to use for the momentary speed. In a wider context, indirect communication via the environment is of high importance in distributed systems [6, 2].

Stigmergy is usually considered as the indirect communication between agents via the environment without targeting a specific recipient. Stigmergy has been used to explain nest building, sorting, and foraging in social insects [13, 3, 2].

To our knowledge this work is the first information-theoretic characterization of stigmergy. It allows us to refine the understanding of stigmergic communication by including the *relationship* between agents and the environment (cf. [1]). For example, as a reminder to do something one could go to sleep in one of several rooms. Later, using the room one wakes up in as a cue, one could engage in different corresponding activities. Information may thus be communicated via one's relation with the environment.

This paper is structured as follows. Sec. 2 introduces the relevant concepts of information theory. Sec. 3 presents the model of the perception-action loop we are using. Sec. 4 describes our experimental approach. In Sec. 5 we evolve an agent to stigmergically communicate with itself through time. In Sec. 6 we evolve pairs of agents to engage in classical stigmergic communication. The work is summarized and discussed in Sec. 7.

2 Information Theory

We denote random variables with uppercase letters, e.g., X, their sets of values with calligraphic letters, e.g., $\mathcal{X}$, and the values with lowercase letters, e.g., x. In this paper we deal exclusively with discrete variables. By abuse of notation we denote the probability that $X = x$ as $p(x)$ and the conditional probability of $X = x$ given that $Y = y$ with $p(x|y)$.

The *entropy* of X, denoted by $H(X)$, is defined as a measure of uncertainty of the probability distribution of X:

$$H(X) := -\sum_{x\in\mathcal{X}} p(x)\log_2 p(x)$$

The *conditional entropy* of X given Y, denoted $H(X|Y)$, is defined as uncertainty of X knowing Y weighted by the probability of a particular realization of Y occurring:

$$\begin{aligned} H(X|Y) &:= \sum_{y\in\mathcal{Y}} p(y)H(X|Y=y) \\ &= -\sum_{y\in\mathcal{Y}} p(y)\sum_{x\in\mathcal{X}} p(x|y)\log_2 p(x|y) \end{aligned}$$

The *mutual information* between X and Y, denoted $I(X;Y)$, is defined as reduction in the uncertainty of X given Y:

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

All of the above quantities are *nonnegative*. We always calculate them using the binary logarithm, hence they are measured in *bits*. An important property of information-theoretic measures is that they do not depend on the particular values of the variables – *the measures only depend on the probability distributions* of the values. For a detailed introduction to information theory consult, for example, [4].

3 A Model of the Perception-Action Loop

In this section we present a model of the perception-action loop of an agent. The constituents of the loop, which are modeled as random variables, are sensors S, actuators A, the memory M of the controller, and R – the rest of the world. We need R to formally account for the effects of actuation and the environment on the sensors.

As in [15] we interpret the perception-action loop in terms of a communication channel-like model. In order to model the temporal aspects we *unroll* the loop in time by introducing a time variable t. To account for the complete loop we model the dynamics of arbitrary number of time steps, as opposed to [15] where only one time step is modeled.

We model the relations between the variables as a causal Bayesian network [11] which is a directed acyclic graph where any node, given its parents, is conditionally independent from any other node which is not its parent or successor (any node directly or indirectly reachable from the node). In our model this property results in conditional independence from the past.

We show the pattern of relations between variables at two consecutive time steps on Fig. 2. We assume that the pattern

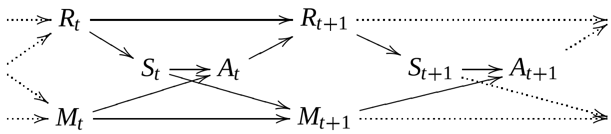

Figure 2: A temporal section of the unrolled perception-action loop modeled a Bayesian network. The dotted lines show how the section connects to adjacent sections. S – sensoric input, A – action performed, M – memory of the controller, R – the rest of the world.

of relations is time-invariant and thus holds for any t. Thus, the graph on Fig. 2 is just a section of the network.

The diagram can be read as follows: A_t is picked given S_t and M_t. Furthermore, S_t is obtained from R_t.

The perception-action loop is created by information flowing from sensors to actuators, and from actuators via the rest of the world to future sensoric input. If these flows extend over more than one time step, they may be mediated by a combination of internal and external variables potentially enabling complex behavior, such as filtering based on the state of the environment or the memory.

4 Experimental Approach

We study simple stigmergic scenarios and track information flows through the environment and the perception-action loop. We measure the flows using information-theoretic tools. We evolve agent controllers to maximize various information flows passing via the environment.

4.1 The Testbed

We base our experiments on a model which despite its simplicity captures important features of systems employing gradient sensors.

The environment consists of a two-dimensional grid of infinite size. A source is located at the center of the grid. The source emits a signal, the strength P of which in any cell of the grid is $P(d) = d^{-2}$ ($P(0) = 2$), where d is the distance to the source. The exact relation is not important for our experiments – it is only important that the decrease is strictly monotonic with distance.

An agent is situated in a single cell at a time. The agent has a gradient *sensor*. The gradient points to the cell with highest signal strength among the four adjacent cells (north, east, south, west). If there are several cells with highest signal strength, the gradient randomly points to one of these with equal probability (see Fig. 3). The agent also has an *actuator* – at each time step the agent performs one of the four available actions: move north, east, south or west.

Following the model of perception-action loop presented in Sec. 3 and its notation, we denote the sensoric input (the gradient) with the random variable S with values in set $\mathcal{S} = \{s_N, s_E, s_S, s_W\}$; the action with the random variable A

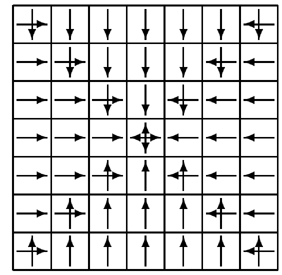

Figure 3: Sensoric input S vs. the position R. In cells with only one arrow the input is constant. In cells with multiple arrows it is randomly chosen from the listed options. The source is located at the center of the grid.

with values in set $\mathcal{A} = \{a_N, a_E, a_S, a_W\}$; the internal state or the memory of the agent's controller with the random variable M with values in $\mathcal{M} = \{1, 2, \ldots, N\}$, where N is the number of states; the two-dimensional position of the agent with the random variable R with values in $\mathbb{Z}^2$. S, A, M, and R completely describe the state of the world at any instant.

According to the model in Sec. 3 A_t and M_{t+1} depend directly only on (S_t, M_t). The agent's controller can thus be seen as implementing the mapping: $(S_t, M_t) \mapsto (A_t, M_{t+1})$. We assume that the mapping is time-invariant.

The mapping can be implemented by a finite-state automaton operating with input set $\mathcal{S}$, output set $\mathcal{A}$, and state set $\mathcal{M}$. Here we use deterministic automata hence allowing only for deterministic mappings. However, without any loss of generality our approach can be used with stochastic mappings implemented by nondeterministic finite-state automata. Determinism of the controller is an experimental choice, not a limitation of the model.

To summarize, the world is an infinite two-dimensional grid. A symmetric signal field is created around a source located at the center of the grid. The strength of the field decreases strictly monotonically with distance. The environment is static. An agent moves on the grid, one cell at a time, in one of the four adjacent cells (north, east, south, west). The agent has a nondeterministic sensor which can distinguish between four directions of the local signal gradient. The agent is controlled by a controller, which is modeled as a deterministic finite-state automaton taking current sensoric input and producing an action. All the controller has access to is its own memory and the momentary sensoric input.

4.2 Information Flow

In the experiments presented in this paper we use a special case of information flow. We "inject" information independent of the past and present state of the system into a variable X (e.g., M_0) by making its distribution independent and entropic. The *information flow* from X to any variable Y in the system is then $I(X;Y)$. This gives us a characterization of stigmergy if X is from one agent and Y is from another.

In general, it is possible to measure information flow without injecting independent information. However, this is out of scope of this paper.

4.3 Measuring Information Flows

We use Monte Carlo simulations for estimating the information flows. Accordingly, to spot and avoid undersampling, for each of the possible initial states of the system we produce 32, 256, or more samples depending on the particular quantities measured. We then increase the number of samples by at least a factor of 16 to check whether the quantities of interest remain stable.

4.4 Evolution

We use evolution as a search in the space of controllers. A minimal setup is used. This is to emphasize that nothing in our general approach is specific to the particular model or the search tool employed. To evaluate the fitness of a controller, it is allowed to control the agent in the particular setup of the experiment. The fitness is expressed as the information flow specific to the experiment.

We initialize the population with five randomly generated controllers. In every generation five best controllers are selected into the next generation and also produce five offspring each. Thus the size of the population is between 5 and 30.

An offspring is produced from its parent by mutation. To speed up the search and make it more efficient we have incorporated ideas from simulated annealing and tabu search: (1) the number of mutations performed is uniformly distributed between 1 and $1+(G \bmod 20)$, where G is the generation; and (2) we do not add offspring controllers which have been evaluated before or are already present in the population. On our problems these adjustments do improve the efficiency of the search. Additionally, we perform at least five separate evolutionary runs to sample different solutions.

The transition matrix of a controller is represented by a mapping $\mathcal{S} \times \mathcal{M} \to \mathcal{A} \times \mathcal{M}$. This mapping can represent any finite-state automaton. We limit our search to deterministic automata only, hence the mapping can be represented as an array of length $|\mathcal{S}| \cdot |\mathcal{M}|$ with each element containing a value between 1 and $|\mathcal{A}| \cdot |\mathcal{M}|$ corresponding to the action to perform and the next state to go into. A mutation is performed by setting a randomly chosen element of the array to a randomly chosen value in the range.

5 Stigmergy for One

Here we study a special case of stigmergy where an agent has to pass information to itself through the environment over a fixed number of time steps. We initialize the agent with information, let the agent run for a while, then erase its memory and let the agent reacquire the lost information.[1]

[1]The source code of experiments is available on request.

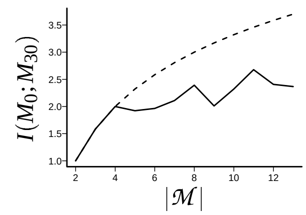

Figure 4: Fitness of best evolved controllers. Solid line – fitness measured as the information flow $I(M_0;M_{30})$ for controllers with $|\mathcal{M}|$ states. Dashed line – theoretical upper bound $\log_2|\mathcal{M}|$.

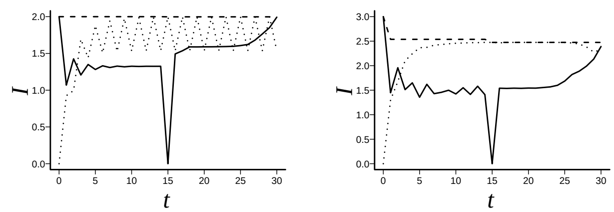

Figure 5: Information flows constructed by the best evolved 4-state (left) and 8-state (right) controller. Solid line – $I(M_0;M_t)$, dotted – $I(M_0;R_t)$, dashed – $I(M_0;M_t,R_t)$.

The agent starts at the center of the grid. We inject $\log_2|\mathcal{M}|$ bits of information into its controller's internal state M_0 by making its probability distribution uniform. At time step 15 we erase the controller's memory by setting it into state 1. We then measure how much of the information "injected" into M_0 is contained in M_{30}. In terms of information flow we want to find a controller which maximizes the information flow from M_0 to M_{30}. As we erase M_{15}, at time step 15 the flow can pass only via the environment. The amount of flow is measured as $I(M_0;M_{30})$.

5.1 Results and Discussion

We have evolved separate populations of controllers with 2 to 13 states for 1000 generations. Evolution does indeed find controllers capable of *self-stigmergy* (Fig. 4).

To analyze the behavior of evolved controllers we track how the injected information "diffuses" through the memory and the position of the agent by measuring $I(M_0;M_t)$, $I(M_0;R_t)$, and $I(M_0;M_t,R_t)$ (Fig. 5).

The environment used in this experiment is static. The agent offloads information into its own position R, that is into own *relation* with the environment. This phenomenon can be explained by the fact that the environment, as perceived by the controller, is dynamic. The evolved controllers employ this to offload and reacquire information.

6 Stigmergy for Two

In this section we study a classical case of stigmergy where one agent (the *sender*) indirectly communicates with another agent (the *receiver*) by changing the environment.

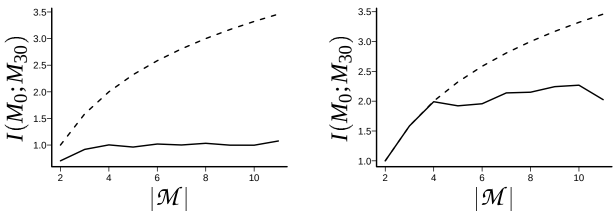

Figure 6: Fitness of best evolved controllers. Left: initial position of the sender is distributed over 11×11 cells, Right: the sender always starts at the source. Dashed line – theoretical upper bound $\log_2 |\mathcal{M}|$.

To enable the environment to be modified we introduce a pushable box. We denote the position of the box with a random variable B with values in $\mathbb{Z}^2$. When an agent hits the box it is moved one cell in the direction the agent last moved. For example, if the agent hits the box having moved north, the box is moved north one cell. There is *no special sensor for the box.* To facilitate the perception of the box with minimal changes to the model, we make the box emit a signal same way as the source does. Signals from the source and from the box are summed. Therefore, the gradient field depends on the position of the box and in principle enables an agent to capture some information about the position of the box using the gradient sensor.

The sender starts with its initial position R_0 distributed uniformly over a square of 11×11 cells centered at the source. We inject $\log_2 |\mathcal{M}|$ bits of information into its controller's internal state M_0. At time step 15 we remove the sender and introduce the receiver by placing it at the source. The receiver's controller is set to state 1. At time step 30 we measure how much of the information about the sender's memory M_0 is contained in the receiver's memory M_{30}. As the receiver's initial state and position are independent of those of the sender, the information flow from the sender's memory into the recipient's memory at time step 15 can pass only via the position of the box B_{15}.

We use evolution to find *pairs* of controllers maximizing the flow $I(M_0; M_{30})$. Instead of individual controllers the evolutionary algorithm operates on sender-receiver pairs.

6.1 Results and Discussion

We have evolved pairs of controllers with 2 to 11 states for 1000 generations. At most 1 bit of information is stigmergically communicated (Fig. 6, left). More information gets communicated if we perform the same experiment but with the sender always starting at the source (Fig. 6, right).

The information flow from the sender's memory M_0 to the receiver's memory M_{30} at time step 15 goes exclusively via the position of the box B_{15}. Thus the amount of information about M_0 in B_{15} is the maximum the receiver could in principle recover. The flow through the memory ($I(M_0; M_t)$) and the position of the box ($I(M_0, B_t)$) is shown on Fig. 7.

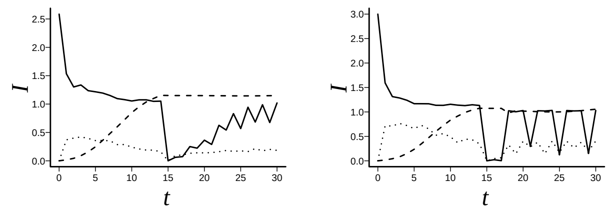

Figure 7: Information flows constructed by the best evolved 6-state (left) and 8-state (right) controller pairs. Solid line – $I(M_0; M_t)$, dashed – $I(M_0; B_t)$, dotted – $I(R_0; M_t)$.

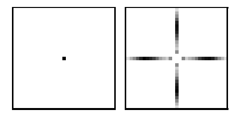

Figure 8: Implicit representations of states in M_0 by box position B_{15} as used by most controllers. Each frame shows a mapping from an initial state onto B_{15}. Intensity of cells reflects the probability of finding the box there: white – zero, black – highest.

Occasionally the receiver reacquires some information about the sender's initial position R_0. The corresponding information flow from R_0 into memory M_t is visualized on Fig. 7. This finding parallels the fact that one's activity usually leaves traces which might be perceived and exploited by others, potentially resulting in a race between strategies for leaving fewer traces and for discovering them.

To understand how the injected information is represented in the position of the box B_{15} we visualize the mapping $M_0 \mapsto B_{15}$ (Fig. 8). It turns out that virtually all evolved controllers use the same representation with minor differences: depending on M_0 the box either remains in the center of the grid, or is pushed away in one of the four straight lines heading north, east, west or south.

7 Conclusions

We have presented an information-theoretic approach to quantifying information flows in agent-environment interactions. In addition to quantification, the approach allows us to "inject" a piece of information into the system and then track how the information diffuses. An advantage of the information-theoretic quantities is that they ignore representation and deal only with the underlying information. This creates a versatile, powerful and flexible view where information is treated as a measurable commodity.

To show the approach in action we have evolved agent controllers for stigmergic behavior. We have characterized stigmergy as the offloading and reacquisition of Shannon information in the relationship between agents and the environment. Stigmergy thus naturally lends itself to being treated in terms of Shannon information flow between agents. As a testbed we have used a two-dimensional grid world where agents have access to a gradient sensor only. The agents are controlled by finite-state automata controllers with limited amount of memory. Although the model is very simple we believe it pertains to a range of Artificial Life models, especially those concerned with stigmergy.

In the first experiment an agent was provided with information which it had to offload from and later reacquire back into its memory, resulting in a kind of self-stigmergy. At first sight a surprising finding is that the agent was able to perform this task without modifying the environment in the common sense. The environment was static and the agent offloaded information *into its own position* in the environment. In other words, the information was offloaded into the *relation* between the agent and the environment.

In the second experiment one agent (the sender) was provided with information to stigmergically communicate to another agent (the receiver) *by modifying the environment*. Interestingly, in addition to the required information the sender offloaded some extra information about itself, some of which was later occasionally acquired by the recipient. This parallels the fact that one's activity usually leaves traces which might be perceived and exploited by others, potentially resulting in a race between strategies for leaving fewer traces and for discovering them.

The sender could only communicate indirectly by pushing a box. The communicated information was thus stored in the position of the box. Neither the sender nor the receiver had explicit sensors for the position or the proximity of the box. The agents did not evolve any "concept" of the box to stigmergically communicate either, rather they managed to create a suitable information flow by interacting with the environment.[2] This emphasizes the fact that the information-theoretic approach enables us to avoid imposing our own biases or assumptions on the agents and thus potentially allows more efficient solutions to be found.

We use the information-theoretic approach based on tracking information flows in order to understand stigmergy and ultimately more general phenomena in a *quantitative* manner. There is strong evidence that interactions of biological agents are partly due to the need for acquiring, exchanging and concealing information. Therefore it is important that the approach enables us to measure and also construct various information flows without much bias, without imposing our own models on the interactions. We believe this may provide novel insights into adaptive systems.

Acknowledgments

We would like to thank the Condor Team from the University of Wisconsin, whose High Throughput Computing system Condor (`http://www.cs.wisc.edu/condor`) enabled us to conveniently run large numbers of simulations on ordinary workstations.

[2] This approach is consistent with the *enactive* view [16] that considers agents as creating the conditions for their interactions.

References

[1] Mark A. Bedau. Dynamics of the environment for adaptation in static resource models. In J. Kelemen and P. Sośik, editors, *Advances in Artificial Life*, pages 76–85. Springer-Verlag, 2001.

[2] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press Inc., 1999.

[3] Eric Bonabeau, Guy Theraulaz, Jean-Louis Deneubourg, Serge Aron, and Scott Camazine. Self-organization in social insects. *Trends in Ecology & Evolution*, 12(5):188–193, May 1997.

[4] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

[5] Kerstin Dautenhahn, Daniel Polani, and Thomas Uthmann. Guest editors' introduction: Special issue on sensor evolution. *Artificial Life*, 7(2):95–98, 2001.

[6] James Hollan, Edwin Hutchins, and David Kirsh. Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, 7(2):174–196, Jun 2000.

[7] Edwin Hutchins. How a cockpit remembers its speeds. *Cognitive Science*, 19(3):265–288, 1995.

[8] David Kirsh and Paul Maglio. On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4):513–549, 1994.

[9] Chrystopher L. Nehaniv. Meaning for observers and agents. In *IEEE International Symposium on Intelligent Control / Intelligent Systems and Semiotics, ISIC/ISAS'99*, pages 435–440. IEEE Press, 1999.

[10] Stefano Nolfi and Davide Marocco. Active perception: A sensorimotor account of object categorization. In B. Hallam, D. Floreano, J. Hallam, G. Hayes, and J-A. Meyer, editors, *From Animals to Animats 7*, pages 266–271. MIT Press, 2002.

[11] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2001.

[12] Daniel Polani, J. T. Kim, and T. Martinetz. An information-theoretic approach for the quantification of relevance. In J. Kelemen and P. Sośik, editors, *Advances in Artificial Life*, LNCS, pages 704–713. Springer, 2001.

[13] Guy Theraulaz and Eric Bonabeau. Modelling the collective building of complex architectures in social insects with lattice swarms. *Journal of Theoretical Biology*, 177(4):381–400, Dec 1995.

[14] Hugo Touchette and Seth Lloyd. Information-theoretic limits of control. *Phys. Rev. Lett.*, 84:1156, 2000.

[15] Hugo Touchette and Seth Lloyd. Information-theoretic approach to the study of control systems. *Physica A*, 331(1–2):140–172, January 2004.

[16] Francisco J. Varela, Evan Thompson, and Eleanor Rosch. *The Embodied Mind: Cognitive Science and Human Experience*. The MIT Press, 1993.