Contingent robot behaviors from self-referential dynamical systems

Ralf Der, Frank Hesse and René Liebscher
Universität Leipzig, Institut für Informatik
POB 920
D-04009 Leipzig
{der|fhesse|liebschr}@informatik.uni-leipzig.de

March 11, 2005

Abstract

The self-organization of behavior is both a striking phenomenon in living beings and a challenging objective for autonomous robots. In our earlier work we introduced homeokinesis – the dynamical pendant of homeostasis – as a general domain invariant principle for behavioral self-creation. The present paper continues these investigations under a more pragmatic aspect. We start from the formulation of two requirements to the behavior namely that actions are such that (i) they are maximally sensitive reactions to the sensor values and that (ii) the consequences of the actions taken are still predictable. We show how this general statement can be formulated into a concrete error function E measuring the distance between the current and the ideal behavior formulated by the requirements. Gradient descending E produces a self-regulating dynamical system. Mathematical arguments show that the robot behaviors emerging from this are both explorative and sensitive to the environment. From the general principle simple learning rules are derived for the neurons of a closed loop robot controller. These learning rules are shown in a simple application with a physical robot to realize a self-learning autonomous robot which can survive in a sufficiently simple world without any further external help. In particular we demonstrate that sensors are automatically integrated according to their response strength as soon as they deliver a signal to the controller. Moreover the system also can deal with the problem of a rapid change in the properties of the sensors.

1 Introduction

The central interest of our work is the self-organized acquisition of behaviors for autonomous robots. Our work is based on the belief that true autonomy must involve the phenomenon of emergence. Before giving some ideas how this could be realized in the robotic domain let us first illustrate the goal in a realistic case. Consider a robot with a neural network controller with synaptic weights initially in the *tabula rasa* condition. So there is no reaction of the robot to its sensor values and activities, if present at all, are only stochastic ones. The robot is to be in an environment with static and possibly also dynamic objects. The task now is to find an objective function for the adaptation of the controller which is entirely internal to the robot driving the parameters so that the robot will start acting and while acting to explore and develop its perception of the world and of object related behavior.

Our work aims at finding general principles for the realization of this program. In our earlier work we started from some basic system theoretic ideas which find their roots in the old cybernetics tradition combined with more recent knowledge from the theory of dynamical systems. A brief reference to these and other related work will be given below. In the present paper we give a reformulation of this principle from a more pragmatic perspective. The principle to be put forward in the present paper is based on two considerations. On the one hand we want the robot to adapt to its environment. This is done by endowing the robot with a minimum of cognition in terms of a world model which allows the robot to predict the consequences of its actions in a narrow time horizon at least. In the present work this will be realized simply by a parameterized function approximator. Then to behave in a predictable way is our first requirement to the behavior of the robot. On the other hand we require that the robot chooses its actions as the most sensible reaction to its sensor values. It is this latter point which brings the robot to activity as will be seen explicitly below. The behavior is emerging as a compromise between these two opposing tendencies.

One of the results of the present paper consists in showing that this extremely simple and general principle can be formulated into a concrete objective for the adaptation of the robot. Moreover it will also be seen that the parameter dynamics (learning rules) for the robot derived from this principle are both rather simple and biologically plausible. Yet the emerging behaviors are by far not trivial. Before presenting these findings in Sec. 3 we will in Sec. 2.1 present our closed-loop control paradigm by considering an elementary SM loop. This serves the purpose of introducing the dynamical system description of the sensorimotor loop and discussing special effects of the noise and in particular introduce the concept of hysteretic control which will play a central role throughout the paper.

The principle can be translated into concrete learning rules for the synapses of the controller neurons. In the present paper we want to investigate these learning rules in a simple case and show that they generate an explorative behavior of the robot which is highly sensitive to the reactions from the environment (by the frequency effect, see Sec. 4.2.2 below). Control realized on the basis of the general principle is achieved in tight sensor motor coupling meaning that all sensors responding to the motor activities are automatically integrated into the generation of the motor command. This will be presented in the subsequent sections.

In concluding this introduction we will give a few remarks on related work. Homeostasis as introduced by Cannon [Cannon, 1939] and later Ashby [Ashby, 1954] is a basic incentive when looking for general principles explaining the functionality of complex self-organizing biological systems. The principle of homeostasis has recently received new attention in neurosciences. In particular in a series of papers by Turrigiano and others, cf. [Turrigiano and Nelson, 2000] have considered the role of various homeostatic mechanisms serving the purpose of counterbalancing the destabilizing effect of Hebbian learning. In [Turrigiano and Nelson, 2004] the idea is that there are specific homeostatic plasticity mechanisms that dynamically adjust synaptic strengths in such a way that the stability of the network is sustained, see also [Feldman, 2002]. There are a few attempts to introduce homeostatic mechanisms in robotics, cf. [Paolo, 2003], [Williams, 2004] but so far the benefit of these approaches is not really established.

While obviously helpful in stabilizing systems the principle of homeostasis seems not well suited for the construction of behavior systems. In fact the aim of such a system is not stasis but a common kinetic regime shared by the constituents of the system in order to produce the behavior in the world. We introduced homeokinesis [Der et al., 2002], [Der et al., 1999], [Der, 2001] as the dynamical pendant of homeostasis.

The behaviors emerging from such general principles are contingent. This means that they depend strongly on the specific initial and environmental conditions and on the specific physics of the robot. In this way our work is also a contribution to the fostering and further understanding of the role of embodiment in the creation of artificial beings, see Refs. [Pfeifer and Scheier, 1999], [Lichtensteiger and Pfeifer, 2002] and others. The focusing on behavior as arising from an entirely internal perspective is also an objective of the constructivistic approach [von Glasersfeld, 1995] and of autopoiesis, cf. [Maturana and Varela, 1979] which underlines the internal perspective of the agent. Our contribution to these developments is to provide a concrete mathematically grounded approach for the realization of these ideas in real robots.

2 Closed-loop control in the sensorimotor loop

Throughout the paper we consider velocity control of a mobile robot under the closed loop control paradigm. This means that the velocity of the wheels is directly derived from the sensor values, in particular those of the wheel counters itself. In order to discuss the peculiarities of this approach let us consider at first a very simple system.

2.1 An elementary sensorimotor loop

Let us consider a closed-loop velocity control of a robot with the sensorimotor loop closed via the wheels alone. For the sake of simplicity we consider the one-dimensional case, i.e. the robot can move only along a straight line. Our controller consists of a single leaky-integrator neuron under the rate coding paradigm.

2.1.1 Dynamics of the loop

The membrane potential z is updated in the time step $t \dots t+1$ as

$$\tau \Delta z_t = -z_t + c x_{t+1} + H \tag{1}$$

where the input is the true wheel velocity x as measured by the wheel counter and H being a bias (threshold). The update is carried out when the new sensor value x_{t+1} arrives. The output of the neuron

$$y_t = \tanh\left(z_t\right) \tag{2}$$

is the target wheel velocity of the robot. The true wheel velocity x_{t+1} as read back by the wheel counter may be assumed to

$$x_{t+1} = a y_t + \xi_t \tag{3}$$

where the parameter a is a hardware constant and ξ incorporates all effects due to slip, friction, discretization noise and so on which make the true velocity deviate from the model assumption $x_{t+1} = a y_t$. The constant a

(the response strength of the channel) can be learned by minimizing the error

$$E = (x_{t+1} - a y_t)^2 \tag{4}$$

for samples (x_{t+1}, y_t) obtained on-line in each time step $t = 0, 1, \ldots$.

Using eq. 1 we find the closed dynamical system

$$\tau \Delta z_t = -z_t + ca \tanh(z_t) + H + c\xi_t \tag{5}$$

describing the time evolution of the membrane potential of the robot. Together with eq. 2 this dynamics completely defines the behavior of the robot.

The concrete behavior obviously depends essentially on the values of c and H. The essential features of this dynamics are illustrated best by using an alternative view of the dynamics of eq. 5 obtained by considering the update of z as a gradient descent on a potential V

$$\tau \Delta z_t = -\frac{\partial}{\partial z_t} V(z_t) + c\xi_t \tag{6}$$

Using

$$\frac{\partial}{\partial z} \ln(\cosh(z)) = \tanh(z) \tag{7}$$

we find that

$$V(z) = -R \ln(\cosh(z)) + \frac{z^2}{2} + Hz \tag{8}$$

with $R = ca$ being the feed-back strength in the sensorimotor loop. As usual the gradient dynamics of eq. 6 may be visualized by that of a sphere sliding down on the walls of a vessel filled with a viscous fluid, see Fig. 1. In our case we may use the small z approximation $\tanh z = z - z^3/3$ to get the more simple expression

$$V(z) = \frac{1}{2}(1-R)z^2 + \frac{1}{12} R z^4 - Hz \tag{9}$$

for the potential which is well known from many branches of physics and dynamical systems theory.

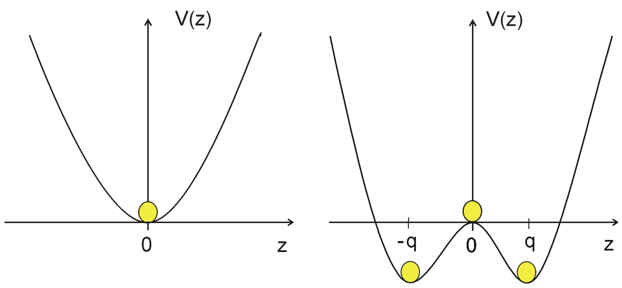

Figure 1: The state dynamics of the sensorimotor loop can be considered as a gradient descent on the potential $V(z)$, assuming $H = 0$ here. Maxima of the potential correspond to instable fixed points, minima to stable ones. The figure shows the potential with $R < 1$ (left) and $R > 1$ (right) where it is a double well potential with two stable and one instable fixed point. Each sphere represents a fixed point of the system.

The fixed points of the system correspond to the extrema of the potential, the stable fixed points being at the minima of the potential. In Fig. 1 (left) the potential is plotted for $R < 1$ where the fixed point at $z^* = 0$

is indicated by the sphere. With increasing R we have a pitchfork bifurcation at $R = 1$ and for $R > 1$ we get a bistable system, i.e. we have one instable fixed point at $z^* = 0$ and two stable fixed points at

$$z^* = \pm \sqrt[2]{\frac{3(R-1)}{R}} \tag{10}$$

as indicated by the spheres in the double well potential in Fig. 1 (right). The deviation of this approximate value from that for the exact potential of eq. 8 is in the region of a few percent in the pertinent (see below) region $1 < R < 1.3$.

2.1.2 Peculiarities of closed-loop control

As described above, the velocity is not given by some external description but is leveling itself as a result of the dynamics of eq. 5. In particular if $R > 1$ we may initialize the robot with any starting velocity and after some time its velocity will approach one of the two possible fixed point values, i.e. the robot will move either forward or backward with constant velocity. Which of the fixed points is realized depends on the starting value and possibly the noise.

One of the benefits of this closed loop control system (under the assumption $R > R_c$, H being 0) consists in the following.

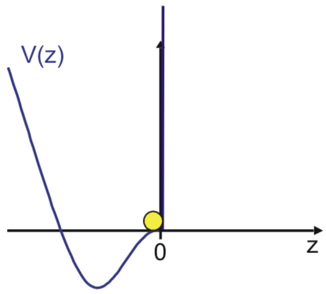

Figure 2: The collision with an obstacle in the potential picture of the dynamics. The robot was moving forward, i.e. the state was at the r.h.s ($z > 0$) minimum of the double well potential. The impenetrable object corresponds to an infinitely steep rise in the potential so that the state is bound to move to the left hand minimum and the robot starts moving backward.

When colliding with an obstacle the wheels are blocked, so that $x = 0$ and z decays. If there is some (small) additional noise in the dynamics of the membrane potential z it will fluctuate around zero. These fluctuations can be amplified if they are of the right sign, i.e. if the robot is moving away form the obstacle. Hence after a short time the robot is found to move away from the obstacle, cf. Figs. 2 and 3. We may say that in this elementary sense the robot is able to survive. This has been corroborated nicely in practice with different kinds of robots.

2.1.3 The role of the bias. Hysteresis.

In the case of finite H the fixed points and hence the velocity of the robot are obtained from

$$z = R \tanh(z) + H$$

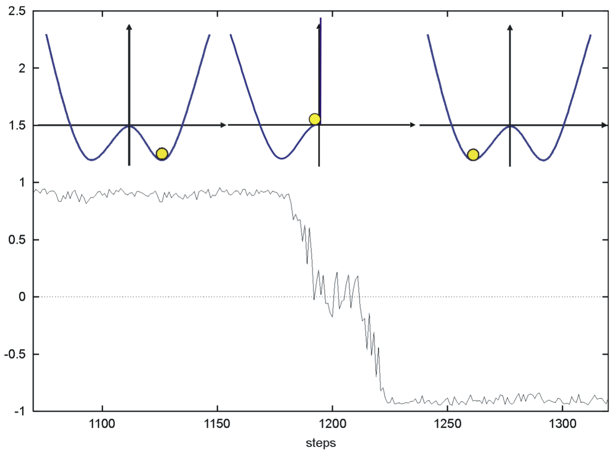

Figure 3: Neuron output y of the closed loop control system (with feed-back strength $R > R_c$) before and after a collision with an obstacle obtained from an experiment with a real Khepera robot (bottom). Above the neuron output curve the corresponding fixed points of the system are represented by a sphere on the potential $V(z)$ (top). Before the collision the system is in the fixed point with positive sign (robot drives forward). During step 1185 to 1215 the robot is kept at the instable fixed point $z = 0$. Around step 1215 the robot starts moving backward because of the noise amplification effect.

so that they also depend on the value of H and there is a hysteresis effect w.r.t the change of H for the case that $R > 1$. The hysteresis effect results from the fact that there is a region $-H_c < H < H_c$ of bistability. Outside we have only a single FP which has the same sign as H. When moving the value of H from outside into the region of bistability the FP realized depends on which side of the outer region one is coming from, see the Fig. 4 for details. The dependence of the system from both parameters H and R can be seen in Fig 5. With $R < 1$ there is only one fixed point $z = 0$, but with $R > 1$ we see the hysteresis effect in the (z,H) space which is the larger the larger R.

2.1.4 Quasi-equilibrium

In many cases of practical interest one may assume that the dynamics of the membrane potential z is fast as compared to the changes in the sensor values so that Δz is small in eq. 1 and we may approximately write

$$z_t = cx_t + H$$

so that the controller output y is a direct function of the sensor values. Under this assumption we get a closed update rule for the sensor values as

$$x_{t+1} = ag\left(cx_t + H\right) + \xi_t$$

By the same token we may also write

$$z_{t+1} = cag\left(z_t\right) + H + c\xi_t$$

We will work with these approximations below.

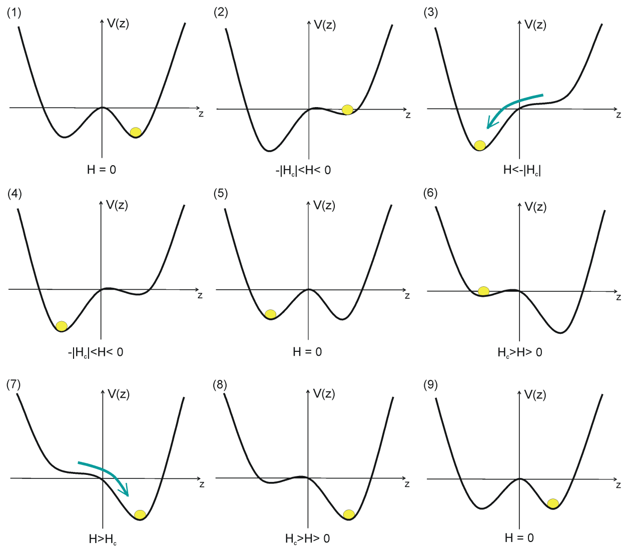

Figure 4: The hysteresis cycle. The diagrams show the stages of one hysteresis cycle starting from $H = 0$ (diagram (1)) with the state at $z > 0$ as represented by the sphere. Decreasing H leads to a deepening of the left minimum, while the right minimum gets more flat, but the state remains at the minimum at $z > 0$, see diagram (2). If $H = -H_c$ both the maximum at $z = 0$ and the right minimum disappear so that the system shifts to the left minimum of the potential (3). Increasing H until $H = 0$ brings us back to the initial situation with the difference that the system changed to the fixed point with negative sign, cf. diagram (4,5). The diagrams (6) and (7) show the switching from the minima at $z < 0$ to the minima at $z > 0$ by increasing H. By decreasing H until $H = 0$ the hysteresis cycle is finished, see diagram (8,9).

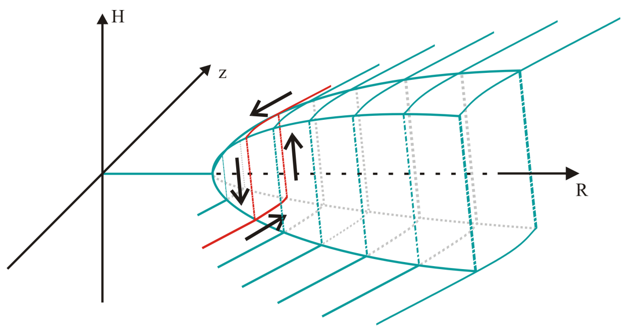

Figure 5: State and parameter dynamics of the system for adiabatic changes. With $R < 1$ there is only one fixed point $z = 0$, but with $R > 1$ we see the hysteresis effect in the (z, H) space which is the larger the larger R.

2.2 Sensorimotor loop in the general case

In the general case we have a vector of sensor values $x_t \in R^n$ at the discrete instants of time $t = 0, 1, 2, ...$ By way of example we may consider the Khepera robot where

$$x = (v_l, v_r, IR_1, ..., IR_8)^T \tag{11}$$

with v_l and v_r are the wheel velocities of the left and right wheel, respectively, as measured by the wheel counters, IR_i is the value of the infrared sensor i with $0 \leq IR_i \leq 1$. Closed loop control means that the controller is given by a function $K : R^n \to R^m$ mapping sensor values $x \in R^n$ to motor values $y \in R^m$

$$y = K(x)$$

In the example we have $y = (y_1, y_2)^T$, y_i being the target velocity of wheel i. Our controller is to be adaptive, i.e. it depends on a set of parameters $c \in R^C$.

2.2.1 World model and sensorimotor dynamics

We assume that our robot has a minimum ability for cognition. This is realized by a world model $F : R^n \times R^m \to R^n$ mapping the actions y of the robot on the new sensor values, i.e.

$$x_{t+1} = F(x_t, y_t) + \xi_t \tag{12}$$

where ξ_t is the model error. The model F can be learned by the robot using any learning algorithm of supervised learning. Let the model be a parameterized function (neural net) with parameters $a \in R^M$. The parameters a can be adapted by gradient descending the error function based on ξ. The structure of the model and the learning procedure define the passive cognitive abilities of the robot.

With these notions we may write the dynamics of the sensorimotor loop in the closed form

$$x_{t+1} = \psi(x_t) + \xi_t \tag{13}$$

where

$$\psi(x_t) = F(x_t, K(x_t))$$

The function ψ can be visualized as a time series predictor for the time series of the sensor values x_t.

By way of example we may consider the elementary SM loop of Sec. 2.1 where now

$$K(x) = g(cx + H)$$

with the controller parameters c and H (the parameter (bias) H in the example is always considered explicitly because its role is more that of an additional internal state of the neuron)

$$F(x, y) = ay$$

and

$$\psi(x) = ag(cx + H)$$

so that the sensorimotor dynamics is

$$x_{t+1} = ag(cx_t + H) + \xi_t$$

2.2.2 The dynamical systems approach to robot control

The time discrete stochastical dynamic system eq. 13 is a mathematical description of the sensorimotor dynamics. Our approach is based on the dynamical systems formulation and tends to adapt the controller so that the robot behavior which is the manifestation of the dynamical system has the desired properties. Using the dynamical system as a substrate for the robot behavior have been considered by several authors under varying contexts and with varying success. Related to our subject is the work by Jun Tani, cf. [Tani, 2004], [Tani and Ito, 2003] and of people around Gregor Schoener, [G. Schöner and Engels, 1995], [Steinhage, 1997], [Hock et al., 2003] An elaborate behavior based design system has been developed in the context of dual dynamics. The system has a layered structure of behavioral subsystems realized by ordinary differential equations, each layer having its own time constant. Interactions between the subsystems is realized by specific interaction and "bifurcation-inducing" mechanisms which have to be designed by hand, cf. [Bredenfeld et al., 2001].

The authors quoted have mainly tried to design dynamical systems such that they realize prescribed tasks, the smooth navigation through a cluttered environment being a prominent example. The main difference with our paper is that we design an objective for the self-regulation of the dynamical system without a concrete task given from outside. The behaviors emerging are therefore contingent but the interesting point is that in the interplay between destabilizing the sensorimotor dynamics and staying nevertheless predictable is the route towards the emergence of environment related behavior.

3 Principles of self-regulation

We are now going to find an objective for the adaptation of our robot based on entirely intrinsic principles. Before doing so we formulate the common approach to adaptive systems.

3.1 Adaptive systems

The dynamics depends on the parameters c of the controller K. Changing c changes the behavior of the system. A system is adaptive if there is an objective function measuring the distance from the current to a desired behavior. (This might be as abstract as measuring the survival properties of the system). One realization of the adaptation is a parameter dynamics as gradient flow

$$\Delta c_t = -\varepsilon S\left(x_t, c_t\right)$$

where

$$S = \frac{\partial}{\partial c} E$$

so that we have the combined dynamics

$$\begin{aligned} x_{t+1} &= \psi\left(x_t, c_t\right) + \xi_t \\ c_{t+1} &= c_t - \varepsilon S\left(x_t, c_t\right) \end{aligned} \tag{14}$$

The problem of learning in robotics consists in finding this objective E for the generation of a desired behavior. In the usual approach to adaptive system the function E is provided from outside and is to be designed such that the system so to say establishes the desired reference to the environment. We will consider so called self-referential systems, i.e. systems for which the objective function is derived from the dynamics of the system itself. We will present in the following one such objective in an explicit way. The aim of the present work is to study these systems and derive some general properties of their dynamics.

3.2 Being sensitive but predictable – the paradigm of controlled acuteness

Of course there are many different paradigms to define a self-referential system. One idea is to reduce the influence of the noise, i.e. the unpredictable part of the behavior. Then the predictability of the future would be the central objective of the adaptation of the behavior. However it is clear that the best realization of this paradigm is given with a "do nothing" behavior. If cognition is to understand what happens, then cognition with this kind of behavior is optimal (in a static world at least). In order to get a robot with an internal drive for activity we require in addition that the reactions of the robot to its sensor values are qualified by a maximum sensitivity.

We claim that one can combine the two requirements by introducing as a first step virtual sensor values $\hat{x}$ defined by

$$\|x_{t+1} - \psi(\hat{x}_t)\| = \min \tag{15}$$

with a conveniently defined norm. Explicitly the shift $v = v_t = \hat{x}_t - x_t$ is

$$v_t = \arg\min_u \|x_{t+1} - \psi(x_t + u)\| \tag{16}$$

The shift is a measure of the sensitivity of the function ψ towards changes in its arguments. In fact, using $x_{t+1} = \psi(x_t) + \xi$ in eq. 16 we immediately see that the shift necessary to achieve the change ξ in the value (output) of the function ψ is the smaller the more sensitively ψ reacts to changes in its arguments (inputs). This will be discussed on a more formal basis in Sec. 3.3 below.

Obviously v is small if both ξ (which measures the predictability) is small and the function ψ is sensitive. Hence the two aims of getting a robot with both highly sensitive reactions and predictability of behavior amounts to the requirement that the shift necessary to produce the new sensor values is as small as possible. We may consequently define

$$E = \|v\|^2 \tag{17}$$

as our objective function measuring the deviation between the current and the desired behavior. We use the Euclidean norm

$$\|v\|^2 = v^T v = Tr\ vv^T$$

where Tr is the trace of a matrix.

Gradient descending E yields to the parameter dynamics

$$\Delta p = -\varepsilon \frac{\partial E}{\partial p}(x, c) \tag{18}$$

where p is any of the parameters (controller and world model) on which ψ depends. Note that the state dynamics eq. 13 and the parameter dynamics eq. 18 run concomitantly. Taking into account only the parameters c of the controller (assuming the world model being learned) we have a combined dynamics in the space $\mathbf{R}^n \times \mathbf{R}^c$, i.e. we have to consider the dynamical system

$$\begin{aligned} x_{t+1} &= \psi\left(x_t, c_t\right) + \xi_t \\ c_{t+1} &= c_t - \varepsilon \frac{\partial E}{\partial c_t}\left(x_t, c_t\right) \end{aligned} \tag{19}$$

This system may be called self-referential since the parameter dynamics is induced by the system dynamics as represented by ψ (i.e. by the current knowledge of the system dynamics) alone.

3.3 Explicit expressions

The definition eq. 15 of the shift may be written as the requirement

$$\psi\left(x\right) + \xi = \psi\left(x + v\right) \tag{20}$$

If v is small we may use Taylor expansion to write

$$\psi\left(x + v\right) = \psi\left(x\right) + L\left(x\right) v \tag{21}$$

where the Jacobian matrix L is defined as

$$L_{ij} = \frac{\partial}{\partial x_j}\psi_i\left(x\right)$$

Using eq. 21 in eq. 20 we find

$$v = L^{-1}\left(x\right)\xi$$

and obtaining v means now "only" to find the inverse of the matrix L provided the latter exists.

Eq. 17 may now be written

$$E = \left\|L^{-1}\right\|_D^2$$

where we introduced the weighted matrix norm

$$\left\|A\right\|_D^2 = Tr\ \left(A^T A D\right)$$

with

$$D_{ij} = \xi_i \xi_j$$

being the matrix of correlations between the model errors (noise) in the different sensoric channels.

The explicit expression displays the main properties of the gradient flow in the parameter space induced by the gradient descent on E. On the one hand E will be small if the vector $\xi = x_{t+1} - \psi\left(x_t\right)$ is small, i.e. if the robot behaves in the most predictable way. On the other hand the Jacobian matrix determines the local stability of the dynamical system defined by ψ. With L in the denominator of E the gradient descent will destabilize the sensorimotor dynamics. This is tantamount to an increase in the sensitivity of the system towards its sensor values since the Jacobian measures also the response strength of the function ψ with respect to changes in its arguments. From these simple arguments it can already be anticipated that bestowing the sensorimotor dynamics with a gradient flow in parameter space driven by E will produce a system with very rich properties.

4 Parameter dynamics in the elementary sensorimotor loop

Now let us return to the simple case presented in Sec. 2.1. We have $x \in R^1$ and $F(x, y) = ay$ so that

$$\psi(x) = ag(z)$$

with $z = cx + H$ (displaying the parameter H explicitly again). The Jacobian is

$$L = \psi'(x) = Rg'(z)$$

with $R = ca$ being the feed-back strength in the sensorimotor loop.

4.1 The parameter dynamics

The error E boils down to

$$E = \frac{\xi^2}{L^2}$$

The derivative is written as

$$\frac{\partial E}{\partial p} = -2\frac{E}{L}\frac{\partial L}{\partial p} - \frac{2}{L}\xi\frac{\partial \psi}{\partial p}$$

where $p \in \{c, H\}$. We assume in the present paper that $\overline{\xi} = 0$ so that the ξ term does not contribute to the parameter dynamics in the average over the noise. Using $L > 0$ (see below) and $g'' = -2gg'$ in the case $g(z) = \tanh(z)$ together with $Rg(z) = z - H$ at the fixed point we obtain

$$\begin{aligned} \Delta c &= \mu a - 2\mu x (z - H) \\ \Delta H &= -2\mu (z - H) \end{aligned} \tag{22}$$

where

$$\mu = 2\varepsilon E/R$$

(R will be seen to be positive below) is a modified update (learning) rate, and $g' = \tanh'(z) = 1 - \tanh^2(z)$. We will see below that the system goes into a limit cycle in the x, H space. Averaging over a period and using that the amplitude of H is much smaller than that of z we may simplify the parameter dynamics further to

$$\begin{aligned} \Delta c &= \mu a - 2\mu xz \\ \Delta H &= -2\mu z \end{aligned} \tag{23}$$

The parameter dynamics is to be used concomitantly with the z dynamics so that the parameters c and H in eqs. 1 or 5 are now time dependent. As we will see below the time scale for the change of in particular H is on the level of the behavior so that in other words the behavior is essentially controlled by the dynamics of H. This is different from the usual paradigm of learning where we have a learning and a performance phase or where there is a separation of time scales for learning and behaving. It may be of interest that behavior control by the synaptic dynamics has been obtained also in the framework of evolutionary robotics, cf. [Nolfi and Floreano, 2000].

4.2 Properties

We have been using the parameter dynamics given by eq. 23 in simulations and with real robots for quite some time. Of the many interesting properties we have encountered in these experiments we will discuss the following.

4.2.1 No bias

Let us consider the case $H = 0$ first. The dynamics for c, eq. 23, consists of the anti-Hebbian term given by the product of the input into the synapse times the membrane potential of the neuron both quantities being felt directly at the synapse. The driving term μa is given by the response strength a of the sensor x to the output of the controller. This term can also be obtained empirically by modulating the neuron output with a periodic perturbation and filtering this signal from the sensor values, hence we may say that the learning rule is a purely local one.

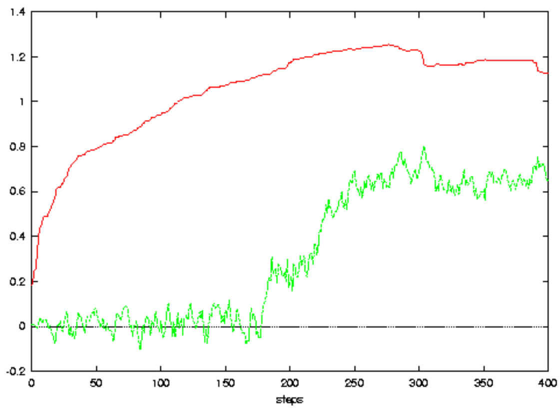

Figure 6: The increase of the feed-back strength R (upper curve) due to the learning procedure. Initially the rise is very steep due to the fact that L is in the denominator. If R is above the critical value the target wheel velocity (lower curve) increases with the sign (direction of motion) determined by the noise amplification effect.

In order to discuss the effects of the two terms we assume that we start the system with $R < 1$ (in the tabula rasa condition, i.e. $0 < R << 1$, e.g.) so that z fluctuates around zero, i.e. the robot executes a random walk. With $z \approx 0$ the anti-Hebbian term is negligible and the driving term is seen to increase the value of R since $\Delta R = \Delta\,(ca) = \mu a^2$, see Fig. 6. Once $R > 1$ is reached the velocity increases exponentially so that the robot starts moving. With $y^2 > 0$ the anti-Hebbian term comes into play and the increase of c is stopped if $a = 2zx$ is reached, i.e. if $1 = 2z \tanh z$ or $1 = 2Ry^2$ which happens at $R \approx 1.2$ corresponding to $y \approx \pm 0.65$. The direction of the robot (sign of the velocity) is arising from a spontaneous breaking of the $x \to -x$ symmetry inherent in the complete (i.e. parameter and state) dynamics. Note that R is the feed-back strength in the sensorimotor loop so that we observe a self-regulation of the system to a feed-back strength which is slightly supercritical.

An interesting phenomenon is observed if the robot hits an obstacle. We have seen above that with R fixed the robot will invert the velocity after some time. In the present case this is accompanied by an increase in the feed-back strength R due to the fact that with $x = 0$ it is only the driving term which is active in the parameter dynamics. Increasing R means increasing the noise amplification effect so that the robot starts being more active at the wall contact, see Fig. 7 for details.

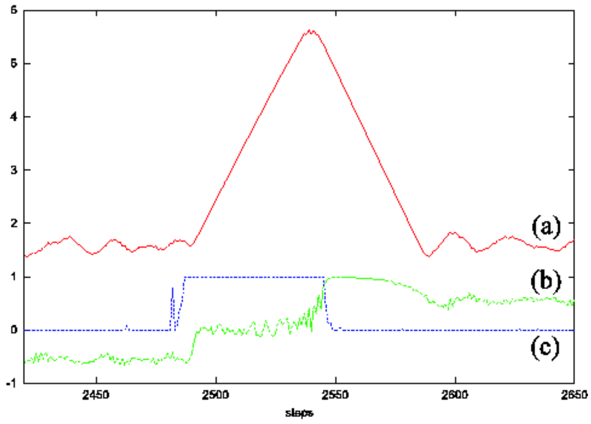

Figure 7: Time course of the synaptic strength c (a) during wall contact of the robot as indicated by the infrared sensors (c). c and hence the feed-back strength is seen to increase during the contact and to decrease again when moving away form the obstacle due to the anti-Hebbian term in the parameter dynamics. The increase in the feed-back strength increases the noise amplification and by this the escape probability, see the y curve (b).

4.2.2 Self-induced search

With $H \neq 0$ there is a hysteresis effect w.r.t the change of H as discussed above. Eq. 23 shows that the change of H always aims at destabilizing the actual fixed point of the system ($\Delta H \sim -z$). Hence the system executes the hysteresis cycle shown in Fig. 4. But the system does not exactly reach the fixed points because of the rapid change of H (as compared to the other parameters). Thus the shape in Fig. 5 is washed-out and it is easily seen that a smooth limit cycle behavior is obtained. The value of c is seen to slightly oscillate with twice the H frequency, but in the average the strength $\overline{R}$ of R is self-regulating again to the slightly supercritical value of $\overline{R} \approx 1.2$, cf. Fig. 8. We may consider the transition to the limit cycle as a self-induced Hopf bifurcation in the x, H space where the value of c is self-regulating to the regime slightly above the bifurcation point.

The frequency of the limit cycle oscillation is modulated by the strength of the noise ξ^2, which we call the frequency effect. With varying noise strength the robot will execute an irregular searching behavior, i.e. the robot will move forward for some time then reverse velocity and move backwards and so on. The most interesting property however is observed when the robot collides with some obstacle so that the wheels get blocked. Then ξ^2 in eq. 22 is very large so that the rate of change of H largely increases and the robot nearly immediately will reverse its velocity, see Fig. 9. In this way the parameter dynamics may be said to create an explorative behavior which stays sensitive to (perturbations by) the environment.

In the applications with both wheels driven by self-regulating neurons the robot is found to explore the

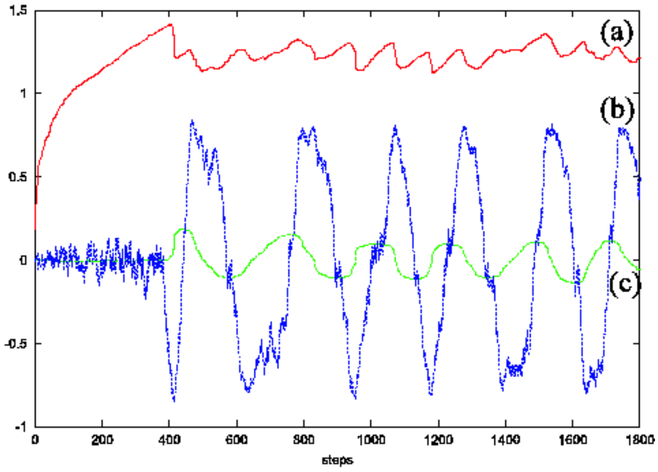

Figure 8: The increase and self-regulation of the feed-back strength R (a) due to the learning procedure. c and therefore R is oscillating with twice the frequency of H (c) which is determined by the strength of the noise ξ^2. The phase shift between H and y (b) is a consequence of the hysteresis effect. At about step 400 the Hopf bifurcation takes place.

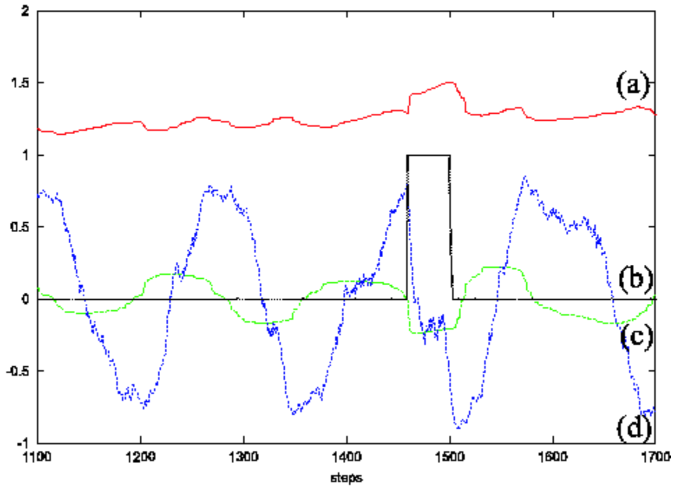

Figure 9: Time course of the response-strength R (a) and bias H (c) during wall contact as indicated by the infrared sensor (b). When the wheels get blocked the model error ξ^2 is very large so that the rate of change of H largely increases and the target wheel velocity y (d) is nearly immediately reversed.

world without getting stuck in corners or at other obstacles, see the videos [Der, 2003]. It should be noted that these properties are not the performance of the trained neuron but instead result from the interplay of state and parameter dynamics, i.e. the concomitant effects of eqs. 1 and 22.

5 Several channels

Let us now consider the case of one controller neuron with several sensoric channels. The sensor values x_i may now depend in a more general form on the value of y. In the present paper we stipulate as in the one-channel case simple proportionality. i.e. we write the sensorimotor loop as

$$x_{i,t+1} = a_i y_t + \xi_t$$

where the deterministic part may be considered as the model and the "noise" ξ is the model error, the constants a_i being learned online. Our aim is to find the properties of the parameter dynamics in this case. The update rule for the membrane potential is

$$\tau \Delta z = -z + \sum_i c_i x_i + H \tag{24}$$

with the fixed point being given by the solution of

$$z = R \tanh(z) + H \tag{25}$$

(we assume $\overline{\xi}_i = 0$ for all channels) where R

$$R = \sum_i c_i a_i$$

is the overall feed-back strength in the sensorimotor loop.

Using as before $y_t = g(z_t)$ we get under the quasi-equilibrium assumption the loop dynamics as

$$x_{t+1} = \psi(x_t) + \xi_t$$

where now $x \in R^n$ and

$$\psi_i(x) = a_i g(cx + H) \tag{26}$$

so that the Jacobian L is

$$L_{ij} = a_i c_j g'(z) \tag{27}$$

Our general principle needs some customizing since the shift v is not uniquely defined. Formally this is seen by the fact that L is not invertible. We may remove the ambiguity by making an assumption on the direction of v. In the present paper we stipulate that v is in the direction of a. The reason behind this is that the inputs x should be produced by the deterministic part of the SM dynamics. Hence x is proportional to a apart from the noise. Using the approximation $\xi = Lv$ and writing $v = u\hat{a}$ (where $\hat{a}^2 = 1$) we obtain

$$\xi = ug'R\hat{a}$$

Considering the projection of ξ on a as the relevant property of the noise we obtain our new objective function as

$$E = u^2$$

where

$$u^2 = \frac{\xi^2 \cos^2 \phi}{\Lambda^2}$$

where ϕ is the angle between the vectors ξ and a and

$$\Lambda = g'(z)\, R$$

Following the lines of Sec. 4.1 we obtain the following rules for the dynamics of the parameters

$$\begin{aligned} \Delta c_i &= \mu a_i - 2\mu z x_i - \gamma \mu c_i \\ \Delta H &= -2\mu z \end{aligned} \tag{28}$$

where $\mu = 2\varepsilon u^2 / R$ and γ (which is small) was introduced in order to produce a (weak) decay of the weights. This is necessary in order to damp that part of the initial conditions which is orthogonal to a. We again find that the change of c_i is given by a driving together with an anti-Hebbian term which can be interpreted biologically.

In order to analyze the parameter dynamics we start with an initializing of c such that the feed-back strength is $0 < R << 1$. Then y is fluctuating about 0 and the driving term in the learning dynamics produces $\Delta c_i = \mu a_i$ so that $\Delta (c_i a_i) = \mu a_i^2$ and hence $\Delta R = \mu a^2$. Obviously the overall feed-back strength R increases with channels of higher response strength $|a_i|$ being favored. Once R exceeds the critical value $R_c = 1$, the anti-Hebbian term comes into play. The parameter dynamics becomes stationary ($H = 0$) if

$$\gamma c_i = \left(1 - 2Ry^2\right) a_i$$

or

$$R = \frac{1}{\gamma / a^2 + 2y^2} = \frac{a^2}{\gamma + 2x^2} \tag{29}$$

so that

$$c_i = \alpha a_i$$

Obviously the c_i reach values so that all sensors are integrated into the sensorimotor loop according to their response strength.

The value of α is obtained as

$$\alpha = \frac{1}{\gamma + 2x^2} = \frac{R}{a^2}$$

Using the fixed point equation in the form $y = \tanh (Ry)$ together with eq. 29 we get

$$y = \tanh \left(\frac{y}{\gamma / a^2 + 2y^2} \right)$$

The position of the fixed point is seen easily to depend smoothly on the value of γ / a^2 for not too large γ. For instance the fixed point is at $y = 0.58$ and $R = 1.15$ if $\gamma / a^2 = 0.2$ which is only slightly lower as compared to $y = 0.65$ and $R = 1.19$ corresponding to $\gamma \to 0$ and also to the one-channel case. The inclusion of the H dynamics will again produce the limit cycle behavior but the preceding results will stay valid in the average over (at least) one period.

Both this simple analysis and computer simulations readily show that the system is self-regulating again into the limit cycle oscillations independently on the number of channels and the values of a_i, since the feedback strength is determined again by R alone. Hence we may say that the irregular, environment-sensitive explorative behavior is reproduced also in the case of many channels where each channel is related to a single sensor. The difference now is that the velocity control of the robot is based on the inputs of all sensors. This in particular means that it will respond to perturbations ξ_i in each channel i in a channel specific manner, i.e. it is most sensitive to channels with a high response factor. This may be of interest in practical applications.

6 Switching sensor activity

One point of interest in the present paper is with the switching activity and characteristics of sensors. In the above considerations we have assumed that the sensor response is essentially proportional to the velocity of the robot. This is not the case for a proximity sensor. In this case we could use a preprocessing and consider the change of the sensor value in the time step as one of our x_i. However the problem is that this sensor characteristics is valid only if the sensor is "on", i.e. the obstacle is in the reach of the sensor. Fig. 10 shows a robotic system in two different situations. In the left part the wheel sensor shows a larger response strength than the infrared sensor and therefore obtains a larger synaptic strength as indicated by the thickness of the arrows from the robot sensors to the controller. The right part depicts a "close to collision" situation where the infrared response is larger than that of the wheel sensor. Another point of interest of the present approach

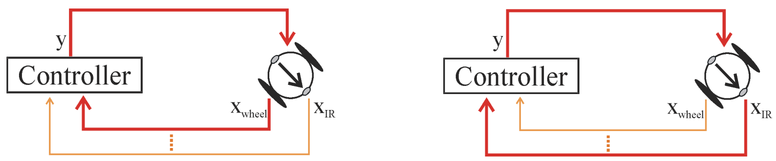

Figure 10: Sensorimotor loop closed over wheel and infrared sensor. If the response factor a of the wheel sensor is larger the synaptic strength of the input from this sensor is greater than from the infrared sensor, as represented by the thickness of the arrows from robot sensors to the controller (left). If in a "close to collision" situation the infrared sensor shows a greater response than the wheel sensor, the synaptic strength corresponding to the infrared sensor will be larger (right).

is the case that new sensors are installed for some time or that sensors temporarily break down. In any case our parameter dynamics will be seen to include the temporary switching on and off of sensors.

In principle the parameter dynamics of eq. 28 may well cope with the situation. Assume we switch on a new sensor k and assume the value of the coupling $c_k = 0$. Then in the beginning we have $\Delta(c_k a_k) = \varepsilon\xi^2 a_k^2 (1 - 2Ry^2) > 0$ since the damping term $-\gamma\mu c_k a_k$ in this channel is negligible as compared to the other channels, because of the small value of c_k. Obviously the value of $c_k a_k$ is rising and this procedure will stop only if $c_k = \alpha a_k$ is reached. Concomitantly the couplings of the other sensors and hence the value of α is readjusted so that the global balance is reestablished. Hence a newly switched on sensor is automatically integrated into the sensorimotor loop. We may also see that the switch-off situation is dealt with in the same automatic fashion.

The problem however is that the processes of readjusting the coupling vector c takes some time. In practical applications (see below) the switching on and off of sensors may take place in very short time intervals. It is therefore of interest not to relearn the couplings but instead to have a kind of long-term memory where the couplings are stored and are read out appropriately. This is possible either on the basis of direct information on the state of the sensors or on context information which is able of qualifying the sensor situation. We will study this case in the following.

7 Experiments with a physical Khepera robot

Our experiments have been chosen to show in a simple case that the derived learning rules generate an explorative behavior of the robot which is highly sensitive to the reactions from the environment. This means that it will move more or less tentatively as long as the knowledge is small, i.e. the modeling error is large (the a_i are still erroneous) and with increasing knowledge a more and more explorative behavior will originate, while it stays sensitive to the environment. Furthermore we want to investigate the situation of switching sensor activities. We therefore consider a Khepera robot in a moveable box which on its hand is confined in a larger area with fixed walls as borders, see Figure 11.

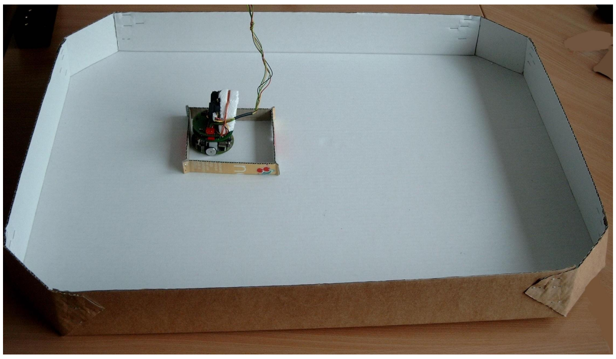

Figure 11: Khepera robot inside a moveable box which is situated in a larger box with fixed walls.

In the experiments two sensors are used, one measuring the velocity of the wheels $x_1(t) = v(t)$ and a pseudo-infrared sensor. The output $x_2(t) = p(t)$ (writing $x(t)$ for x_t and so on for notational reasons) of the pseudo-infrared sensor is proportional to the target velocity y of the robot so that it can be used immediately as one of the x_i in the parameter rule given above. This is done to retain the simple case of linear dependencies for easy understanding, while showing the properties of the approach. However the pseudo-sensor is triggered by the physical infrared sensors with outputs $r(t)$ (the moveable box is large enough to contain both *on* and *off* regions of the pseudo-infrared sensor)

$$p(t) = \begin{cases} by(t-1) & : \quad \max_{i=0}^{7} r_i(t) > r_{min} \\ 0 & : \quad \max_{i=0}^{7} r_i(t) \leq r_{min} \end{cases} \tag{30}$$

so that the sensor $p(t)$ is active as long as there is at least one infrared sensor with an activity larger then the threshold r_{min}, else the value of this sensor will be zero. Thus we have the case of a switching sensor.

The controller of the robot consists of a single neuron with the update rule for the membrane potential as given above. Its output $y \in [-1, 1]$ is the forward velocity of the robot. We used several tricks in order to make the learning more effective. On the one hand the learning of H is much faster then that of the other c since this one realizes the quick reaction of the robot to large model errors and hence the frequency effect. On the other hand with such large learning rates we must take some precautions against divergencies. Therefore we push each update through a squashing function, cf.

$$\Delta c \leftarrow \eta \tanh\left(\frac{1}{\eta}\Delta c\right)$$

so that there is no change in the case of small Δc but there is a maximum step width given by η.

Finally, the model parameters a_1 and a_2 are learned on-line by gradient descending the model error eq. 4. However in the model the measured speed depends linearly on the controller output, i.e. the model is appropriate only if the robot moves without problems. When colliding with the wall the model is not valid any longer and the learning should be switched off. This is achieved in our case by multiplying the learning rate by a kind of reliance factor which is chosen as

$$r_j = \exp\left(-\beta\xi_j^2\right)$$

for channel j where $\xi_j^2 = (x_j - a_j y)^2$ is the error of the model in this channel.

7.1 Long-term memory for the parameters

For many robotic applications the problem of switching sensor activity plays an important role. In particular our pseudo-infrared sensor $p(t)$ can switch frequently between $p(t) = 0$ and $p(t) = by(t)$ (see eq. 30) according to the robots position in the moveable box (Figure 11). So good predictions are obtained only if the model parameter is switched according to the situation.

The sensor situation depends on the output of the physical infrared sensors r_i which can be transformed in a context $m(t)$ with the values

$$m(t) = \begin{cases} 1 & \max_{i=0}^{7} r_i(t) > r_{min} \\ 0 & \max_{i=0}^{7} r_i(t) \le r_{min} \end{cases}$$

Our solution to this problem is to train a neural network (with the context $m(t)$ as input) to set the value of the model parameter $a_2(t)$ according to the context. The learning signals are directly given by gradient descending the model error.

Moreover different model parameters a_i produce different synaptic weights, see the discussion above. Therefore the controller parameters are also represented by a neural network (with the context $m(t)$ as input), except for the bias $H(t)$ which is changing rapidly all the time (compared with the other parameters) and therefore does not need to be memorized.

With the incorporation of this long-term memory our controller is able to handle switching sensors.

7.2 Results

In the experiments the model parameter for the wheel channel $a_1(t)$ was conveniently initialized by hand. The initial value of the model parameter for the pseudo-infrared channel $a_2(t)$ was set to zero and the learning of the model parameter was disabled for the first 1000 steps. As a result the model error is large if the pseudo-infrared sensor is active which leads to an almost immediate change of the value of the bias H so that the robot changes its direction of motion. The effect is that the robot avoids collisions with the walls of the moveable box it is enclosed in.

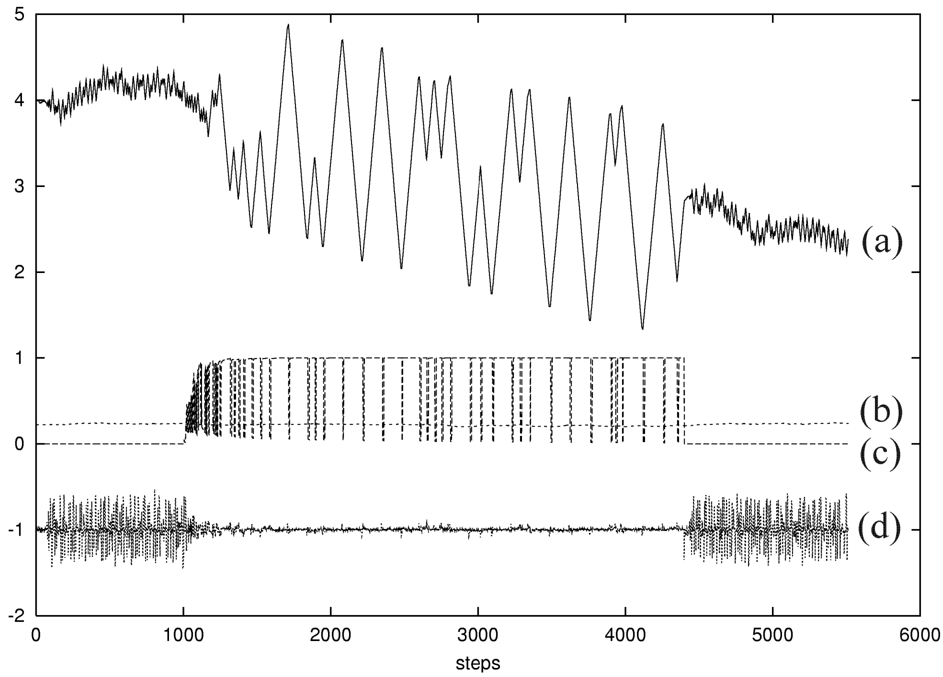

Figure 12: (a) The path traveled by the robot (with odometry error). For the first 1000 steps model learning was disabled so that the robot moves only very cautiously not pushing the moveable box. Then model learning is allowed and with increasing knowledge of the world the robot starts moving the box and explores the full range of the arena while moving the box around. In the end learning was disabled (with a_2 set to 0) again so that the cautious behavior reappears. (b) Model parameter $a_1(t)$ of the wheel channel is already learned at the beginning of the experiment. (c) Model parameter $a_2(t)$ unlearned at the beginning, then learning to jump between 0 and 1 depending on the activity of the infrared sensors and reset to zero at the end. (d) $\xi(t)-1$; The difference between predicted and measured pseudo-infrared sensor values ($\xi(t)$) gets smaller when the model learns to predict (middle) and rises with resetting the model parameter $a_2(t)$ to zero.

When model learning takes place after step 1000 the model parameter $a_2(t)$ becomes more and more adapted (jumping between $a_2 = 0$ and $a_2 = b$, cf. Figure 12 (c)) which decreases the model error. Hence when approaching the wall of the moveable box the relearning of H (Fig. 13 (a)) does not take place and the robot starts moving the box around.

Eventually when the robot reaches the wall of the arena (wheels get blocked) the model error is large so that the rapid relearning of the bias H and hence the velocity reversal takes place at this collision event. In

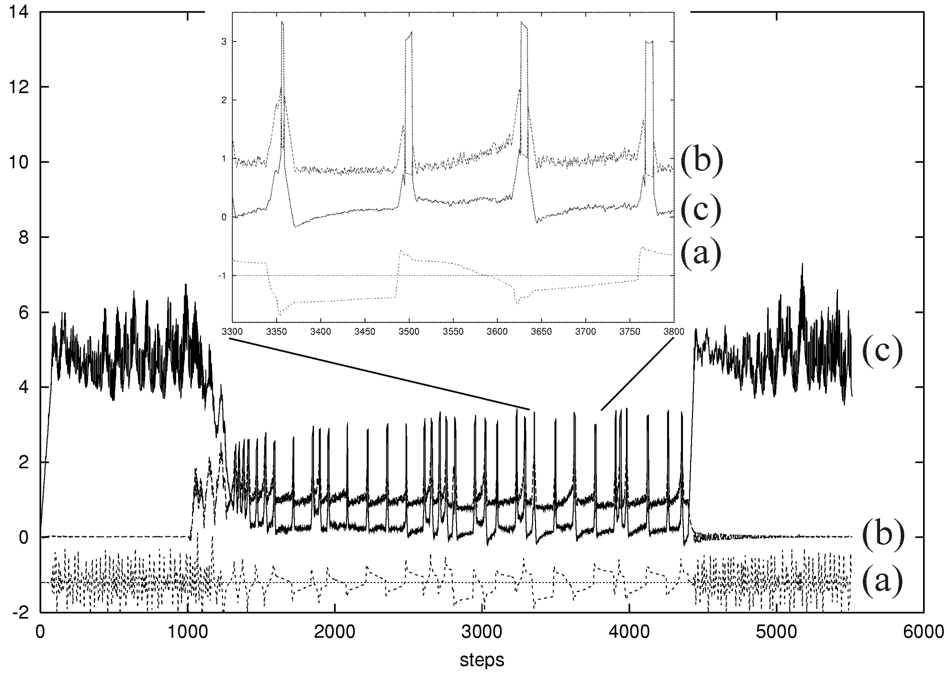

Figure 13: (a) The bias $H(t) - 1$, (b) the weights for the pseudo-infrared input $c_2(t)$ and (c) the wheel input $c_1(t)$ of the homeokinetic neuron used as controller. The input weights converge to different values depending on the value of the model parameter for the pseudo-infrared channel. The frequency of the changes of the bias $H(t)$ and therefore the time the robot travels in one direction depends on the model error.

this way the robot now explores the full region of the arena, see Fig.12 (a).

Approx. with step 4000 the learning is disabled from outside and the model parameter $a_2(t)$ is set to zero again. This leads to a large model error and fast relearning of $H(t)$ when the pseudo-infrared sensor is active. Hence the robot moves only in a short range of his environment like in the beginning of the experiment, so one can see that if the model for some reason is not able to make good predictions any more the robot returns to its uncertain behavior.

As to the controller parameters we observe that at the beginning and at the end of the experiment the pseudo-infrared channel is not included in the sensorimotor loop ($c_2(t) \approx 0$, Fig. 13(b)) because $a_2(t)$ is set to zero. Hence the sensorimotor loop is closed only over the wheels with $c_1(t) \approx 5$ (Figure 13 (c)) and with $a_1(t) \approx 0.23$ this leads to a feed-back strength $R \approx 1.1$ ($\alpha \approx 21.7$).

In the middle part the pseudo-infrared sensor is included in the sensorimotor loop, but only when it is active. Then the model parameter $a_2(t) \approx 1$ leads to the increase of the appropriate weight until $c_2(t) \approx 1$. The readjusted factor $\alpha \approx 1$ can also be seen in the wheel channel ($c_1(t) \approx 0.2$, $a_1(t) \approx 0.23$) so that R is around 1.1 again. When the pseudo-infrared sensor is not active the parameters should be $a_2(t) \approx c_2(t) \approx 0$ and $c_1(t) \approx 5$. This is not the case ($c_2(t) \approx 1$, $c_1(t) \approx 3.3$) so the feed-back strength is smaller then 1 for a short time. The cause is seen in the very fast switching of the pseudo sensor together with the timescale of z which is comparable to the activity time of the sensor. Hence z can not reach the fixed point value, so the theoretically derived values can not exactly be realized.

In a few words this experiment shows the uncertain, tentative behavior (changing direction of motion very often) of the robot in areas the model can not predict properly, the "brave" behavior (covering large areas, which are predictable for the model) and the change between these two behaviors through learning of the model parameters. In doing so the robot stays sensitive to the reactions of the environment. This could be realized through the derived learning rules. Additionally the switching sensor is integrated in the sensorimotor loop, as long as it is active.

8 Outlook and Conclusions

We have demonstrated in the present paper that the very simple learning rules, cf. eq. 23 derived from the general principle realize a self-learning robotic system which can survive in a sufficiently simple world without any further external help. In particular we have seen that sensors are automatically integrated according to their response strength as soon as they deliver a signal to the controller. Moreover the system also can deal with the problem of a rapid change in the properties of the sensors. The basic effect observed is that the learning rule drives the robot into an explorative mode of behavior which however is sensitive to the reactions of the environment by way of the model error. From a dynamic systems point of view we have a closed loop control system with a pitchfork or a Hopf bifurcation (if the learning of the threshold is included) and the effect of the learning is to drive the system to a regime slightly above the bifurcation point where such systems are known to be particularly sensitive.

The findings of the present paper reveal also a general property of our approach which is related to the dilemma between exploration and the exploitation of knowledge gathered in the model. Typically this difficulty arises in the following way. Assume we have an agent which is to explore the world and while exploring to construct a model of its behavior which on its hand is used for the guidance of control. Then one can either choose to stay with the behaviors which are well modeled and be safe or to further explore the space with the chance of getting lost. This needs a careful tuning of the exploration rate in practical realizations.

In our approach this dilemma is solved by the fact that the robot reacts with a change in behavior if the model error is increasing suddenly. This may happen if the robot gets into a situation which is not yet "understood" by its model. This restricts the robot to the behavior which is well modeled by the current model. However with each of such situations the model also gets some new information which can be integrated into the model making the error smaller and hence the reaction slower when the robot reencounters the situation. In this way the robot will conquer increasingly larger regions of the behavior space. In other words our controller automatically adapts the exploration rate to the needs of (model) knowledge acquisition.

Another aspect is in the relation between the short term memory of the original setting where the parameters of the neuron are changed on the behavioral time scale and the long-term memory introduced in the present paper. We have demonstrated that the rapid parameter learning can well be used as a learning signal for a long-term memory which stores the parameter values of the neuron in a context dependent fashion so that a parameter and hence behavior recall guided by context information is realized. In future work this technique will be used for a weighted feed-in of the learning signal into the parameter network the weights being given by a kind of reinforcement signal obtained either from outside or in terms of the future model errors. In the latter case one might use the wheel error so that behaviors which lead to collisions with obstacles would be avoided as a result of the learning. Hence we may expect that the agent acquires more foresight in the course of the time.

Acknowledgement *This work was supported by the DFG under DE798/1–2 in the program 1125 ("RoboCup"). Helpful discussions with Michael Herrmann and Ulrich Steinmetz are gratefully acknowledged.*

References

[Ashby, 1954] Ashby, W. R. (1954). *Design for a Brain.* Chapman and Hill, London.

[Bredenfeld et al., 2001] Bredenfeld, A., Jaeger, H., and Christaller, T. (2001). Mobile robots with dual dynamics. *ERCIM News*, 42.

[Cannon, 1939] Cannon, W. B. (1939). *The Wisdom of the Body.* Norton, New York.

[Der, 2001] Der, R. (2001). Self-organized acquisition of situated behavior. *Theory Biosci.*, 120:179–187.

[Der, 2003] Der, R. (2003). Videos of self-organised robot behavior. http://www.informatik.uni-leipzig.de/ der/Forschung/videos.html.

[Der et al., 2002] Der, R., Herrmann, M., and Molicki, M. (2002). Self-organization in sensor-motor loops by the homeokinetic principle. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, page 510.

[Der et al., 1999] Der, R., Steinmetz, U., and Pasemann, F. (1999). Homeokinesis - a new principle to back up evolution with learning. In *Computational Intelligence for Modelling, Control, and Automation*, volume 55 of *Concurrent Systems Engineering Series*, pages 43–47, Amsterdam. IOS Press.

[Feldman, 2002] Feldman, D. E. (2002). Synapses, scaling and homeostasis in vivo. *Nature Neuroscience*, 5:712 – 714.

[G. Schöner and Engels, 1995] G. Schöner, M. D. and Engels, C. (1995). Dynamics of behavior: Theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems,*, 16:213–245.

[Hock et al., 2003] Hock, H. S., Schöner, G., and Giese, M. A. (2003). The dynamical foundations of motion pattern formation: Stability, selective adaptation, and perceptual continuity. *Perception & Psychophysics*, 65:429–457.

[Lichtensteiger and Pfeifer, 2002] Lichtensteiger, L. and Pfeifer, R. (2002). An optimal sensor morphology improves adaptability of neural network controllers. In *Dorronsoro, José R. (ed.), Artificial neural networks - ICANN 2002. 12th international conference, Madrid, Spain, August 28-30, 2002. Proceedings. Berlin: Springer. Lect. Notes Comput. Sci. 2415, 850-855* .

[Maturana and Varela, 1979] Maturana, H. R. and Varela, F. J. (1979). *Autopoiesis and Cognition: The Realization of the Living.* Reidel, Boston.

[Nolfi and Floreano, 2000] Nolfi, S. and Floreano, D. (2000). *Evolutionary Robotics*. MIT Press,Cambridge, MA.

[Paolo, 2003] Paolo, E. D. (2003). Organismically-inspired robotics: Homeostatic adaptation and natural teleology beyond the closed sensorimotor loop. In Murase, K. and Asakura, T., editors, *Dynamical Systems Approach to Embodiment and Sociality*, pages 19 – 42.

[Pfeifer and Scheier, 1999] Pfeifer, R. and Scheier, C. (1999). *Understanding Intelligence*. Bradford Books.

[Steinhage, 1997] Steinhage, A. (1997). *Dynamical Systems for the Generation of Navigation Behavior*. PhD thesis, Ruhr-Universit at Bochum, Germany.

[Tani, 2004] Tani, J. (2004). Symbols and dynamics in embodied cognition: Revisiting a robot experiment. In Butz, M. V., Sigaud, O., and Gerard, P., editors, *Anticipatory Behavior in Adaptive Learning Systems*, pages 167–178. Springer-Verlag.

[Tani and Ito, 2003] Tani, J. and Ito, M. (2003). Self-organization of behavioral primitives as mulitple attractor dynamics: A robot experiment. *IEEE Transactions of on Systems, Man, and Cybernetics Part A: Systems and Humans*, 33(4):481–488.

[Turrigiano and Nelson, 2000] Turrigiano, G. G. and Nelson, S. (2000). Hebb and homeostasis in neuronal plasticity. *Current Opinion in Neurobiology*, 10:358–364.

[Turrigiano and Nelson, 2004] Turrigiano, G. G. and Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, 5:97–107.

[von Glasersfeld, 1995] von Glasersfeld, E. (1995). *Radical Constructivism - a Way of Knowing and Learning*. Falmer Press, London.

[Williams, 2004] Williams, H. (2004). Homeostatic plasticity in recurrent neural networks. In Schaal, S. and Ispeert, A., editors, *From Animals to Animats: Proceedings of the 8th Intl. Conf. On Simulation of Adaptive Behavior*, volume 8 of *8*, Cambridge MA. MIT Press.