

Learning to feel the physics of a body

Ralf Der¹, Frank Hesse^{1,2} and Georg Martius^{2,3}

¹ University of Leipzig, Institute of Computer Science, POB 920, D-04009 Leipzig

² Bernstein Center for Computational Neuroscience, POB 2853, D-37018 Göttingen

³ Graduate College - Analysis, Geometry and their Interaction with the Natural Sciences

University of Leipzig, Institute of Mathematics, POB 920, D-04009 Leipzig

der@informatik.uni-leipzig.de {georg|frank}@chaos.gwdg.de

Abstract

Despite the tremendous progress in robotic hardware and in both sensorial and computing efficiencies the performance of contemporary autonomous robots is still far below that of simple animals. This has triggered an intensive search for alternative approaches to the control of robots. The present paper exemplifies a general approach to the self-organization of behavior which has been developed and tested in various examples in recent years. We apply this approach to an underactuated "snake" like artifact with a complex physical behavior which is not known to the controller. Due to the weak forces available, the controller so to say has to develop a kind of feeling for the body which is seen to emerge from our approach in a natural way with meandering and rotational collective modes being observed in computer simulation experiments.

1 Introduction

Despite the tremendous progress in robotic hardware and in both sensorial and computing efficiencies the performance of contemporary autonomous robots is still far below that of simple animals. This has triggered an intensive search for alternative approaches to the control of robots. The present paper exemplifies a general approach to the self-organization of behavior which has been developed and tested in various examples in recent years, cf. [7], [2], [3], and [4]. Our objective in the first instance is not the generation of goal oriented behavior but the emergence of sensorimotor coordination in complex robotic systems. In the present paper this is an underactuated snake like artifact with a complex physical behavior which is not known to the controller. Due to the weak forces available, the controller so to say has to develop a kind of feeling for the body which is seen to emerge from our approach in a natural way.

Our paradigm is rooted in the dynamical system approach which we are going to discuss first in order to work out the peculiarities of our approach. Initiated by seminal work about ten years ago, cf. [1], [14] the dynamical systems have found increasing interest in recent years as a substrate for the generation of behavior. The point is that depending on the parameters even low-dimensional dynamical systems offer a broad spectrum of properties ranging from fixed point over limit cycle to chaotic behaviors. By way of example we discuss rhythmic motion generation by limit-cycle dynamical systems. Widely used is a central pattern generator like the one introduced by Matsuoka [10] in 1985 which models the dynamics of two mutually inhibiting neurons in terms of a set of differential equations. This has been further developed and applied to problems like quadruped locomotion [8], drumming [15], juggling [11] playing with a slinky toy [17], and in particular to the gait control of bipedal walking machines which are of much importance in humanoid robotics, cf. [9].

These and related approaches try to use the potentials of dynamical systems in a kind of building blocks approach to complex behaviors. The idea is to use (linear) oscillators in order to realize movement primitives which then are coupled together to a nonlinear dynamical system affording a large range of potential behaviors. These approaches in a clever way use specific properties like entrainment, frequency and phase locking of coupled oscillator systems in order to scale up to systems at the humanoid level.

The annoying task of tuning manually the parameters of the oscillators and the pertinent couplings is avoided by using either reinforcement learning or learning by demonstration, cf. [13], [12] so that for instance in bipedal walking human-like gaits are realized by this building block approach on the basis of dynamical systems.

Nevertheless the physics of any walking machine largely differs from that of the human gait apparatus so that our goal is not to mimic human behavior. Instead it is our aim to first let the machine try to develop its own way of be-

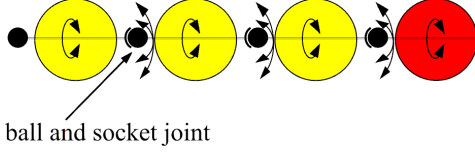


Figure 1. Mechanical construction of the "skidding snake". Ball-and-socket joints between each sphere are drawn. Arrows indicate the possible rotations: in every direction at the joint and around the axis between joints.

ing which afterwards may be sorted and eventually shaped into a desired direction. The corresponding principles of self-organization have been formulated and applied to various robotic systems in our recent papers and will be briefly sketched in Sec. 2 below. Sec. 3 presents the application to an underactuated physical system of many degrees of freedom. Some conclusions may be found in Sec. 4.

2 Self-organizing dynamical systems

Based on the papers [3], [4] we give here the basic principles of our approach. We start from the informations the robot gets by way of its sensor values. Let us consider a robot which produces in each instant $t = 0, 1, 2, \dots$ of time the vector of sensor values $x_t \in R^n$. By way of example we will consider in the present paper the "skidding snake" consisting of a head which can move in the plane driven by forces which may be understood as a force vector $y \in R^2$ of arbitrary direction but restricted magnitude. In a physical realization the head may be a robot with omnidirectional drive. The head is connected to a tail of n ($n = 9$ in the experiment) spheres which are connected to each other by a link (ball-and-socket joint, see Fig. 1) so that each of the spheres can rotate passively like in a string of beads. In the experiments the system is underactuated in the sense that the force acting on the head is not sufficient to drag the tail into arbitrary directions. Instead the motions of the head are more or less lateral only. Nevertheless as we will see in the simulations the controller can excite collective motions by seemingly exploiting inertia and gyro effects.

The sensor values in this example are the velocities of the head relative to the ground, i.e. we measure a vector¹ $x_t = (x_1^t, x_2^t)^T$. The controller is given by a function $K : R^n \rightarrow R^m$ mapping sensor values $x \in R^n$ to motor values $y \in R^m$

$$y = K(x)$$

¹We write the time as an upper index if we want to indicate the components of a vector.

all variables being at time t . In the example we have $y_t = (y_1^t, y_2^t)^T$, y_i^t being the force into the direction $i = 1, 2$ of the plane (in world coordinates).

Our controller is to be adaptive, i.e. it depends on a set of parameters $c \in R^C$. In the cases considered explicitly below the controller is given by the pseudolinear expression

$$K_i(x) = g(z_i) \quad (1)$$

where $g(z) = \tanh(z)$ and

$$z_i = \sum_j c_{ij} x_j + H_i \quad (2)$$

This seems to be overly trivial concerning the set of behaviors which are observed in the experiments. Note however that in our case the behaviors are generated essentially also by an interplay of neuronal and synaptic dynamics (see Eq. 11 below) which makes the system highly nontrivial.

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 $F : R^n \times R^m \rightarrow R^n$ mapping the actions y of the robot on the new sensor values, i.e.

$$x_{t+1} = F(x_t, y_t) + \xi_t \quad (3)$$

where ξ_t is the model error. The model F can be learned by the robot using any learning algorithm of supervised learning.

In the case considered below we have $x, y \in R^2$ and we assume that the response of the sensor is linearly related to the motor command, i.e. we write

$$x_{t+1} = Ay_t + \xi_t \quad (4)$$

where A is a matrix and ξ the modeling error. The model is learned by gradient descent as

$$\Delta A = \varepsilon_M \xi y^T \quad (5)$$

both ξ and y taken at time t . Again, this expression seems to be oversimplified. However model learning will be seen to be very fast so that the model parameters change rapidly in time so that different world situations are modeled by re-learning. Moreover the model only is to represent the coarse response of the world to the actions y of the robot, behavior being organized such that this reaction is more or less predictable. Hence the world model mainly is to give a qualitative measure of these response properties.

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

$$x_{t+1} = \psi(x_t) + \xi_t \quad (6)$$

where in our specific case

$$\psi(x) = AK(x)$$

2.2 Realizing self-organization

As is well known from physics, self-organization results from the compromise between a driving force which amplifies fluctuations and a regulating force which tries to restabilize order in the system. In our sensorimotor dynamics the destabilization is achieved by increasing the sensitivity of the sensoric response induced by the controls based on previous sensor actions. The counteracting forces are obtained from the requirement that the consequences of the actions taken are still predictable. More details may be found in [7], [3], [4]. We formulate these two requirements into an objective function

$$E = \xi^T Q^{-1} \xi \quad (7)$$

where ξ is the model error as introduced above, the positive semidefinite matrix $Q = LL^T$ and L is the Jacobi matrix of the sensorimotor dynamics which in the specific case reads

$$L_{ij}(x) = \sum_{k=1}^n A_{ik} g'(z_k) c_{kj} \quad (8)$$

all quantities depending on time t . Using gradient descent the parameter dynamics is

$$\Delta c_t = -\varepsilon \frac{\partial E_t}{\partial c_t}(x_t, c_t) \quad (9)$$

explicit expressions for the parameter dynamics being given below, cf. Eq. 11. Note that the parameter dynamics Eq. 9 is updated in each time step so that the parameters in practical applications may change on the behavioral time scale. This means that the parameter dynamics is constitutive for the behavior of the robot.

The explicit expression Eq. 7 displays quite obviously the essence of our approach. The matrix Q measures the sensitivity of the sensorimotor loop towards changes in the sensor values. Minimizing E is thus immediately seen to increase this sensitivity since E contains the inverse of Q . We have shown in many practical applications that in this way the robot develops an explorative behavior which however is moderated by the fact that E is also small if the prediction error ξ is small which is the case for smooth environment related behaviors. Behavior may be understood as the compromise between these two opposing tendencies.

3 Experiments

The "skidding snake" was introduced briefly already above. We used the ODE tool (open dynamics engine [16]) for the computer simulations.

3.1 Explicit parameter dynamics

The controller dynamics is given by Eq. 9. The inversion of the 2×2 matrix Q is trivial so that explicit expressions for the parameter dynamics are obtained straightforwardly. In deriving the latter one has to observe that the parameters also enter the g' functions which are treated as

$$\frac{\partial}{\partial c_{ik}} g'_j = \delta_{ij} g''_i \frac{\partial z_i}{\partial c_{ik}} = \delta_{ij} g''_i x_k$$

In the case of $g(z) = \tanh z$ we have $g''(z) = -2gg'$ so that

$$\frac{\partial}{\partial c_{ik}} g'_j = -2\delta_{ij} g'_i y_i x_k \quad (10)$$

The explicit formulae are (omitting the time indices everywhere)

$$\begin{aligned} \varepsilon^{-1} \Delta c_{ij} &= \zeta_i v_j - 2\zeta_i \rho_i y_i x_j \\ \varepsilon^{-1} \Delta H_i &= -2\zeta_i \rho_i y_i \end{aligned} \quad (11)$$

where $\zeta_i = g'_i \mu_i$, $\mu = A^T Q^{-1} \xi$, and $\rho = cv$. We may interpret the first term $\zeta_i v_j$ as a general driving term which is seen to increase the sensitivity of the linearized sensorimotor loop. The second term contains the nonlinearity effects due to Eq. 10 and essentially keeps the neurons out of the saturation regime where they are not sensitive to the inputs. It may be interpreted as an (anti-) Hebbian learning term with strength given by $2\zeta_i \rho_i$, the H dynamics being driven exclusively by this term.

Note that the parameter ε is chosen such that the parameters change at about the same time scale as the behavior. The interplay between synaptic and state dynamics makes the controller a complex dynamic system which is dynamically embedded into the sensorimotor loop. The ensuing robot behaviors thus are of a much larger complexity than the pseudolinear expression with fixed parameters might ever realize.

3.2 Emergence of collective modes

In the experiments the "snake" is moving in an rectangular arena surrounded by walls. There is a considerable friction of the spheres on the ground so that motion essentially requires the spheres to rotate. In our experiment the model and controller parameters (matrices A and C) are initialized as unit matrices. Because of the underactuated setting in the initial phase of the experiment forces applied to the head element lead to small lateral motions of the head with the tail sometimes swaying passively around. Hence the sensor values x and controller outputs y are fluctuating around zero, see Fig. 2. Therefore Δc_{ij} in eq. 11 is dominated by the driving term $\zeta_i v_j$ and the matrix elements of C increase (see Fig 3 top) so that the reaction of the head to

its sensor values becomes more and more sensitive. Hence the amplitude of the controller output increases, but this is more or less only the amplification of the noise and there is no reaction of the body to this random signals. This is seen from the fact that the sensor values keep fluctuating around zero in the first 18000 steps. However with the sensitivity increasing more and more the controller also starts to respond to the swaying of the body and manages to amplify the latter. This process is self-supporting since the emerging collective motions respond to the controller in a much more systematic way. Thus in the long run the controller generates collective motions of all elements which is seen at first as meandering and eventually leads to a rotational mode where the body becomes stiff like a stick due to the gyro effects of the rapidly rotating spheres of the body, see the videos [5], [6].

There is an interesting interplay between the parameter dynamics of the world model and the controller. In the beginning the model parameters are decreasing (first 3000 steps), see Fig. 3 center, since the reaction of the head to the applied forces is weak due to the fact that the controller does not yet "feel" the reactions of the body so that it does not "know" how to excite the collective modes. Around step 18000 the controller starts to excite the collective modes so that there is a definite response of the body to the controller actions which is reflected by the increase of the diagonal elements of the matrix A .

Later on the controller is seen to turn down its reaction strength to the inputs which makes sense since in the rotational mode the forces necessary to keep the rotation going is much smaller. The combined action of C and A is reflected in the response matrix R of the (linear) sensorimotor loop, see Fig. 3 which displays a steep rise in the response of the system when the collective mode is emerging. After a stable period of high rotation frequency the response is seen to decrease and at the end of the rotational period it is seen to have a nice SO2-like form reflecting the fact that the system is in a well controlled rotational mode of lower frequency. Quite generally, the response strength is seen to self-regulate to a slightly supercritical value if the system gives reasonable response.

Another interesting effect is that the frequency of the rotational mode is not constant but increase at first, reaches a plateau value and after some time decreases again. Then the rotational mode decays and reappears after a new fluctuation and meander phase. This play repeats infinitely often which reflects the fact that the system so to say probes into its various modes of behavior. Moreover this remains also true under heavy perturbations by other creatures in the environment, see Fig. 4 and the videos.

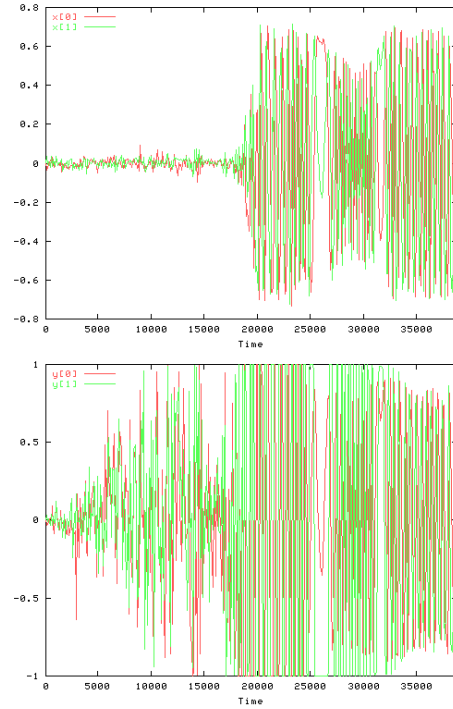


Figure 2. Sensor and motor values during the experiment. Top: Sensor values x fluctuate around zero at the beginning, because the applied forces lead only to small lateral motions of the head with the tail sometimes swaying around passively. With increasing sensitivity the controller starts to respond to the swayings of the body leading to a rotational mode of behavior (second half of the diagram). Bottom: Controller outputs y also fluctuate around zero in the beginning, but with increasing sensitivity of the controller the amplitude increases, the controller responds to the swayings of the body eventually leading to a rotational mode.

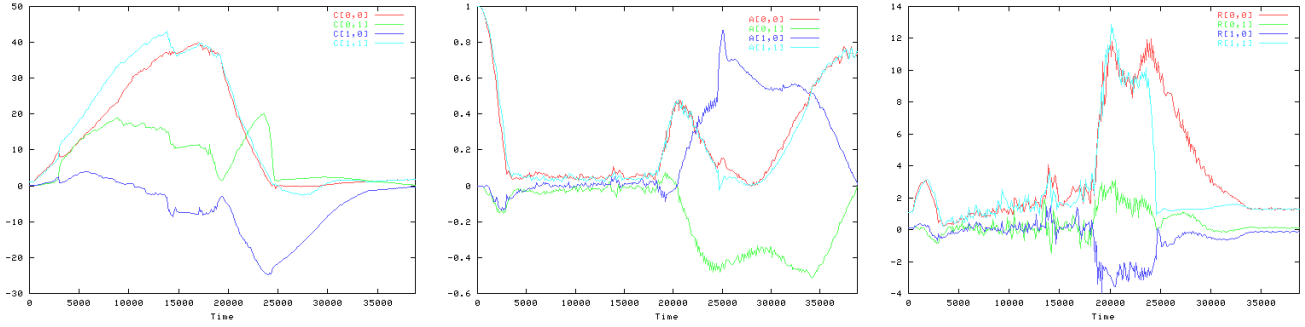


Figure 3. Parameters during the experiment. Left: Matrix C of controller parameters showing the increase of sensitivity, in the first half of the diagram, while in the second half when the collective motion sets in the reaction strength is turned down since there is a definite response of the body to the controller actions even with smaller parameter values. Center: Matrix A of model parameters shows the correlation between sensor and motor values. At the beginning there is a decrease from the initial values to the ones representing the weak reaction of the head element to the controller outputs. After step 18000 when the collective mode sets in the diagonal elements increase since there is a definite response of the body to the controller actions. Right: The response matrix R of the system shows a steep rise when the collective mode emerges. After a stable period the response strength decreases, showing a SO(2)-like form at the end of the experiment reflecting the systems rotational mode of lower frequency. It is seen that the response strength is self-regulating to a slightly supercritical value if the system gives valuable response.

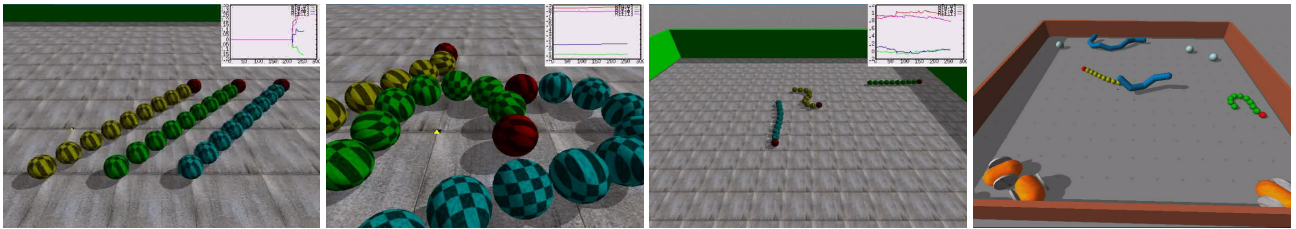


Figure 4. Images of "skidding snakes". Most left: Initial position in an experiment with three "skidding snakes". Left: Close view. Right: In the run. Two "snakes" in rotational mode where the bodies become stiff like sticks due to the gyro effects of the rapidly rotating spheres of the body. The third is in z-shape with the tail passively swaying around. Most right: A group of interacting self-organized creatures controlled by the presented approach (except the three passive balls). The systems try different modes of behavior (including the rotational mode of the "skidding snake") even under these heavy perturbations. See also the videos [6].

4 Discussion

We have demonstrated in the present paper that our general paradigm as formulated mathematically in the error function, cf. Eqs. 7, 8 leading to the parameter dynamics Eq. 11 in applications to a completely unknown body with a complicated physics leads to an active behavior where the controller so to say tries to feel its body and that in the interplay collective modes of the physical system are excited. We have seen that these modes are not persistent but instead are metastable so that the system more or less tries to investigate its behavioral possibilities. Moreover this scenario is stable under strong influences by the environment both of a static and dynamic nature (other agents), see the videos.

An important consequence is also derived for the interplay between the world model and the controller. The system does not have any information on the structure and dynamics of the body so that the world model has to learn this from scratch. This involves the so called cognitive bootstrapping problem meaning that on the one hand the controls are to be such that the world model is provided with the necessary informations. On the other hand these actions require a certain knowledge of the reactions of the body – information is acquired best by informed actions. The concerted manner by which both the controller and the world model evolve during the emergence of the rotational mode seems to be a good example of this process.

We consider our approach as a novel contribution to the self-organization of complex robotic systems. At the present step of our development the behaviors although related to the specific bodies and environments are without goal. As a next step we will realize a so called behavior based reinforcement learning. When watching the behaving system one often observes behavioral sequences which might be helpful in reaching a specific goal. The idea is to endorse these with reinforcements in order to incrementally shape the system into a goal oriented behavior.

Acknowledgment

This work was supported by the DFG under DE798/1–2 in the program 1125 ("RoboCup"). The authors wish to acknowledge the helpful discussions and support provided by Marcel Kretschmann and Naglaa Hamed.

References

- [1] R. D. Beer. A dynamical systems perspective on agent-environment interaction. *Artif. Intell.*, 72(1-2):173–215, 1995.
- [2] R. Der, F. Hesse, and R. Liebscher. Self-organized exploration and automatic sensor integration from the homeokinetic principle. In *Proc. SOAVE Ilmenau 04*, Ilmenau, 2004.
- [3] R. Der, F. Hesse, and R. Liebscher. Contingent robot behavior generated by self-referential dynamical systems. *Autonomous robots*, 2005. submitted.
- [4] R. Der, F. Hesse, and G. Martius. Rocking stumper and jumping snake from a dynamical system approach to artificial life. *J. Adaptive Behavior*, submitted, 2005.
- [5] R. Der, F. Hesse, and G. Martius. Video to the experiments section in this paper. <http://robot.informatik.uni-leipzig.de/Videos/HurlingSnake/2005/CandA+.mpg>, 2005.
- [6] R. Der, F. Hesse, and G. Martius. Videos of self-organized creatures. <http://robot.informatik.uni-leipzig.de/Videos>, 2005.
- [7] R. Der and R. Liebscher. True autonomy from self-organized adaptivity. In *Proc. Workshop Biologically Inspired Robotics*, Bristol, 2002. Internet: <http://www.informatik.uni-leipzig.de/der/Veroeff/bristol.pdf>.
- [8] Y. Fukoka, H. Kimura, and A. Cohen. Adaptive dynamic walking of a quadruped robot on irregular terrain based on biological concepts. *International Journal of Robotics Research*, 22(3-4):187–202, 2003.
- [9] K. Hase and N. Yamazaki. Computational evolution of human bipedal walking by a neuro-musculo-skeletal model. *Artificial Life and Robotics*, 3:133–138, 1999.
- [10] K. Matsuoka. Sustained oscillations generated by mutually inhibiting neurons with adaptation. *Biol. Cybern.*, 52:367–376, 1985.
- [11] S. Miyakoshi, M. Yamakita, and K. Furuta. Juggling control using neural oscillators. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1186–1193, 1994.
- [12] J. Nakanishi and S. Farrell, J. A. and Schaal. Learning composite adaptive control for a class of nonlinear systems. *IEEE International Conference on Robotics and Automation*, pages 2647–2652, 2004.
- [13] J. Nakanishi, J. Morimoto, G. Endo, G. Cheng, S. Schaal, and M. Kawato. Learning from demonstration and adaptation of biped locomotion. *Robotics and Autonomous Systems*, 47:79–91, 2004.
- [14] R. F. Port and T. V. Gelder, editors. *Mind as Motion*. The MIT Press, Cambridge, MA, 1995.
- [15] S. Schaal and S. Kotosaka. Synchronized robot drumming by neural oscillators. In *The International Symposium on Adaptive Motion of Animals and Machines*, Montreal, 2000.
- [16] R. Smith. Open dynamics engine. <http://ode.org/>, 2005.
- [17] M. M. Williamson. Neural control of rhythmic arm movements. *Neural Networks*, 11:1379–1394, 1998.