Summary and outlook

The present thesis is a theoretical contribution to a systematic approach to the self-organization of behavior in autonomous robots. It is deeply rooted in the dynamical system approach to robotics, cognitive science, and artificial intelligence and investigates in some detail the properties of such systems based on a general paradigm. The paradigm tries to give the robot a maximum sensitivity to its sensor values together with a maximum predictability of future sensor values resulting from the actions taken by the robot. Formally, this leads to a parameter dynamics for the neural network controlling the robot which is based on the gradient descent on the so called time loop error, see Sec. 3.3 of Chapter 3. The latter is formulated entirely in terms of the dynamics of the dynamical system describing the robot in its environment (state dynamics). Hence the parameters for the state dynamics are driven by the state dynamics itself and this is why we call them self-referential dynamical systems.

The thesis studies such systems in some detail. The results obtained may be summarized in the following way. A first result is obtained in a very simple sensorimotor loop with delays. In such systems, under the closed loop control paradigm used in the present thesis, one often encounters rapid oscillations of the robot which may be strong enough to destroy the robot. These oscillations have been shown to arise from the time delays and it was demonstrated that by a simple smoothing procedure the problem can be solved. The further chapters of the thesis are devoted to the study of the self-referential systems. On a quite general basis, see Chapter 3, one may argue that the self-referential system tries to maximize its exploration rate in the dimensions of the state space where the model error is large. In this way the robot is enabled to gather more information in the dimensions which are not well covered by the world model, see Sec. 5.3 of Chapter 3. With a good separation of the time scales for the parameter dynamics (learning) and the state dynamics it was shown that the self-referential system generates a fixed point flow of the state variables towards the bifurcation points with occasional jumps to other fixed points, see Sec. 6 of Chapter 3. This is the generalization of the well known situation in the one-dimensional case (with learning of both the bias and the feed-back strength parameters) which was demonstrated in earlier papers to lead to a limit cycle behavior. The thesis adds to these well known results an example for the learning of the world model as driven by the active exploration behavior of the robot in the sense mentioned above, see Sec. 2.2 of Chapter 4.

The main body of results is obtained for the case of a two-dimensional system, see Chapter 5. In the linear case it could be shown that the controller matrix C converges for a large set of initial conditions to an SO (2) structure leading to a periodic or quasiperiodic state dynamics with frequency largely depending on the initial condition of the C matrix. Nonlinearities are shown to largely modify this picture. In order to get a systematics at first the case of learning only the controller matrix C is considered. Several different phenomena could be identified which are realized alternatively depending on the initial conditions and the reaction of the environment. Noteworthy are (i) the case of regular oscillations where a period-4 cycle was shown to be particularly stable; (ii) the case of irregular oscillations with the nondiagonal elements playing the role of a bias dynamics; and (iii) the switching oscillator scenario where one neuron is oscillating and the other one keeps itself at the fixed point, the role of the neurons interchanging after some time. In the latter cases the frequency is seen to be regulated by both the learning rate and the reactions of the environment (noise) so that the behavior of the robot is strongly coupled to the behavior of the world.

The most interesting effects were found for the case that each neuron also has a bias driven as well by the gradient dynamics on the time loop error. The emerging bias dynamics can be considered as an additional internal state dynamics for the controller which increases the complexity of possible behaviors of the robot far beyond a simple reactive behavior. The most prominent effect has been found in form of the so called frequency wandering, which makes the robot to continuously sweep through its accessible periodic behaviors. This is a clear example of its exploration capabilities.

In a practical application, the so called barrel robot, the frequency wandering effect is also realized however in a different way dictated by the embodiment of the agent which is seen to react on the parameter dynamics in a definite way. The results show how the parameter dynamics driving the internal state of the controller sensitively adapts to the specific properties of the body it is controlling. On the phenomenological level one might say that the robot is able to both accelerate and decelerate or that it actively investigates its space of velocities. The sensitive dependence of the combined state-parameter dynamics has also been made explicit by considering the case of noise over the channels, see Sec. 5 of Chapter 5.

In the three dimensional case several of the effects found in two dimensions could be shown to survive in a modified form. In particular, in the case of no bias a theoretical and computer simulation study was carried out showing the persistence of the regular oscillations with higher order periodicity. When the bias dynamics is included, the frequency wandering effect is also observed, however it does not involve all three neurons but instead realizes in subsystems of two neurons, with a switching between the subsystems after some time. The effect is found to be rather stable, i.e. it emerges under a broad range of initial conditions provided the noise is not too strong.

A modification of the general approach has been presented in Chapter 7 by using the so called natural gradient for generating the parameter dynamics from the time loop error. The natural gradient as introduced by Amari into learning theory gives the optimal direction for descending a function if the metric of the space is Riemannian. By using the noise matrix D as the metric tensor we could show that, under the closed loop control paradigm, control is realized on the basis of sensor values weighted according to their feasibility. In the most simple case of uncorrelated noise, this measure is given by the response strength of the sensor divided by the strength of the noise. In more general cases the controls are given by the vector of response strengths a multiplied by the matrix D^{-1}. Moreover an algorithm was proposed and tested in simulations for the on-line learning of this inverse noise matrix.

The present thesis has revealed that the time loop error is a constructive method for the self-organization of a robot behavior which is both explorative and sensitive to the environment. Most of the investigations, however, have been carried through under the assumption that the world (robot + environment) is more or less trivial, i.e. we have proprioceptive sensors only which report on the result of the actions plus some noise. This is reflected in the linear expression $x_{t+1} = Ay_t + \_{t}$ of the sensorimotor dynamics with _ being a pure (and mostly small) random number. An exception has been the barrel robot which is a physical object with large inertia effects and a very complicated relation between motor and sensor values. Nevertheless a strong parallel to the results with the idealized world could be established in the presence of the frequency wandering effect. It would be very interesting to repeat these investigations with different kinds of embodied agents.