

Self-Organization of Exploratory Control

Georg Martius^{*†}, J. Michael Herrmann^{*†‡}

^{*}Bernstein Center for Computational Neuroscience, Bunsenstr. 10, 37073 Göttingen, Germany

[†]Max Planck Institute for Dynamics and Self-Organization, Bunsenstr. 10, 37073 Göttingen, Germany

[‡]IPAB, University of Edinburgh, School of Informatics, 10 Crichton St, Edinburgh, EH8 9AB, U.K.

Email: georg@nld.ds.mpg.de, michael.herrmann@ed.ac.uk

A. Self-organizing control

Animals process sensory information in order to generate behavior that matches the affordances of their environment. The statistical properties of an environment, however, appear substantially different to an inexperienced individual and to its more skilled fellow. On the one hand, appropriate learning opportunities may rarely occur by chance, while on the other hand a concentration on well-learned behaviors tends to reduce the frequency of errors although these are important indicators of the environmental dynamics. In order to extract the currently relevant aspects of the sensory world, the agent should thus aim at performing coherent behavior based on incident regularities as well as maintain a constant excitation of the controlled systems in order to escape from suboptimal behaviors and track the environmental dynamics. The strategies for departing from either of these uninformative situations requires the agent to engage actively in the learning process [1].

In the present contribution we will show that homeokinetic learning [2] is an expedient method to achieve this goal and that it presents a viable approach to active learning. This method comprises the simultaneous maximization of the sensitivity of an agent with respect to its stimuli and the predictability of these stimuli. Sensitization leads to exploratory behavior which is counterbalanced by the requirement of predictability by means of an internal model. As a theoretically provable consequence, the motor activity becomes distributed over many degrees of freedom of a robotic system. Homeokinetic learning leads to flexible, versatile and body-specific behaviors [3] that can be characterized by their controllability.

B. Spherical robots as an exemplary implementation

Among a number of implementations of the learning scheme [3] we focus here on the simulated robot SPHERICAL (Fig. 1b) which is of a relatively simple design, but involves a non-trivial control problem. It is actuated by three weights that are movable inside the robot along orthogonal axes. Any change of the positions of the weights affects the center of mass of the robot and results in a certain rolling movement. Control of this system must account for inertia effects and the intricate relation between motor actions and body movements.

At each time step the controller receives the current positions of the masses as input and calculates new target positions of the three weights on their axes. Simulated motors are used to move the weights to these positions which might be compromised by centrifugal forces. Initially the robot does not

move, but as a consequence of the learning rule the controller becomes more and more sensitive to any changes of its sensor values by amplification of small noisy fluctuations until a more coherent physical movement develops. Later a regular rolling behavior is executed which breaks down infrequently to give way for different movement patterns. Typically, rolling modes around one of the internal axes are seen to occur, see Fig. 1.

C. Exploratory control

The homeokinetic controller produces behavior by self-organization of the sensorimotor dynamics of the robot including physical states and internal parameters. The function of the controller is best understood by considering the map from the sensory state at one time step to the next time step. This map depends both on the actions of the robot and on the environment. The agent has an adaptive internal model that learns to represent an approximation of this map. Those effects of the map that are not captured by the model are considered as errors may they be due to noise, to immaturity, or to a complexity gap between model and environment. Analogously, we can consider the inverse map describing the sensory dynamics in inverse time as well as an corresponding model of this dynamics, which gives rise to a different error. In first order these two errors are related by the Jacobian of the sensory map. The homeokinetic learning rule, in particular, minimizes the backward error and thus simultaneously the inverse Jacobian (maximizing sensitivity), and the forward error (maximizing predictability). This is obviously only feasible when the errors actually have been caused by recent motor actions. Behavioral modes outside the known behavioral manifold correspond to small eigenvalues of the Jacobian, otherwise random fluctuation in these directions had caused corresponding motor actions. This has an interesting consequence for the internal models. If the backward error was indeed determined by a backward model or by a pseudo-inverse of the forward model it would correctly predict small values and had little effect on the learning rule. If, however, the backward error is determined by regularized on-line inversion of the forward model then small sensitivity leads typically to big effects in the learning rule for the controller and hence in the behavior of the robot. The effect depends on the regularization which is to be chosen such that the linear approximation underlying to the approach is not compromised. A further condition is the positivity of the Jacobian. If this is violated fast oscillatory behavior is produced which is, however, typically damped in real robots.

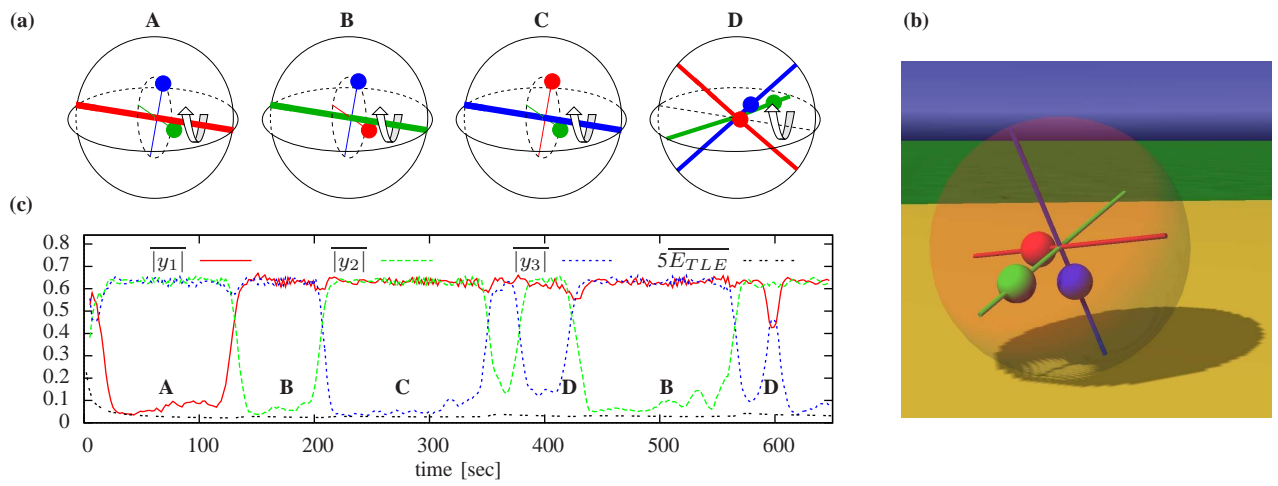


Fig. 1. SPHERICAL robot exploring behavioral modes. The three internal masses are actuated each along its axis. Axes orientations serve as sensors. (a) Typical behaviors: rolling modes about each of the three internal axes (A-C) keeping one weight still, and intermittent rotation about any other (unstable) axis (D); (b) Screen shot taken from computer simulation; (c) Amplitudes of the motor value oscillations ($y_{1...3}$) and the objective function (E_{TLE} (time loop error)) averaged over 10 seconds and scaled for better visibility. An infrequent switching of behaviors is observed and corresponding modes from the panel (a) are labeled. The qualitative behavior is independent of the initial condition.

D. Exploitation of the low-level control structures

As an immediate result of the learning rule, a flexible coordination of movements is observed to appear in more complex robot morphologies [3]. It is also possible to modulate the self-organization process in order to develop specific preferences, cf. [4]. Further applications of homeokinesis include control of myoelectric prostheses, brain-machine interfaces, active signal processing and the interaction of adaptive agents [3].

More interesting is the composition of more complex goal-directed behaviors based on elementary sensorimotor relations that are extracted from the waxing and waning of the emergent behaviors during homeokinetic learning [5]. Typical behaviors for the spherical robot are represented in Fig. 2. Using different phase relations among the actuators the controller is able to stabilize the “natural” behaviors at different speeds. The controllability of such elementary behaviors entails their occurrence and thus their chance to find a representation within the model which in turn improves controllability. In this way, learning becomes a self-stabilizing process that selects certain modes of operation.

On the slower time scale of the learning of the model, however, these modes tend to be destabilized again such that a number of behavior is sequentially activated and learned limited only by the complexity of the internal model. Because the extractable behaviors are negotiated between the dynamics of the robot and the internal model they are well-suited as a set of representative behaviors that can be used in symbolic higher-order learning. Eventually, by a concatenation of several elementary behaviors regions in the environment become reachable that are unlikely to be found by random or quasi-random exploration.

Acknowledgment: We like to thank Ralf Der for very helpful discussions. The project was supported by BCCN Göttingen grant #01GQ0432.

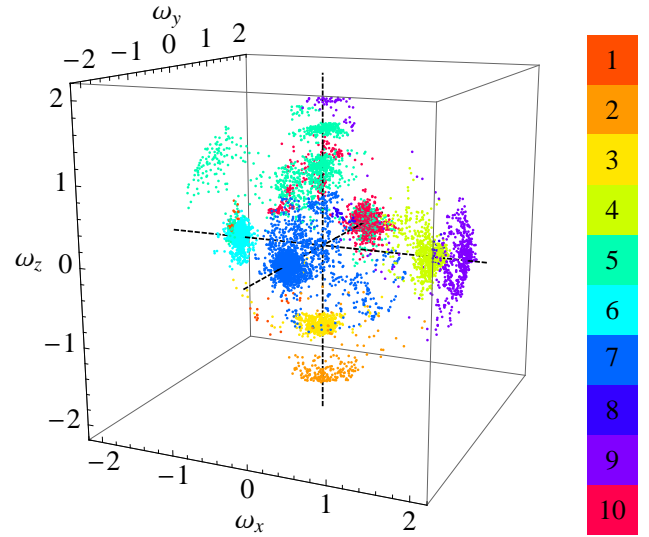


Fig. 2. Partition of the space of angular velocities of the spherical robot by a collective of expert networks. The key on the right assigns expert labels to colors in the picture. The clustering of the data is based on prediction quality achieved by the experts. The angular velocities are the critical variables in the physical dynamics of the robot, they are, however, not directly accessible by the robot, only the axis orientations are sensed and the position of the movable masses are controlled.

REFERENCES

- [1] J. M. Herrmann (2001) Dynamical Systems for predictive control of autonomous robots. *Theory in Biosciences*. **120**:3-4, 241-252.
- [2] R. Der. Self-organized acquisition of situated behavior. *Theory Biosci.*, 120:179-187, 2001
- [3] <http://robot.informatik.uni-leipzig.de>
- [4] R. Der, G. Martius. From motor babbling to purposive actions: Emerging self-exploration in a dynamical systems approach to early robot development. In S. Nolfi et al., eds.: *From Animals to Animats 9*. LNCS 4095, Springer, 406-421, 2006.
- [5] G. Martius, K. Fiedler, and J. M. Herrmann. Structure from behavior in autonomous agents. In *Proc. IEEE International Conference on Intelligent Robots and Systems (IROS 2008)*, 858 – 862, 2008.