Towards controlled self-organized behavioral exploration

Cristina Pinneri and Georg Martius

Max Planck Institute for Intelligent Systems, Tübingen, Germany

October 31, 2017

Introduction

Self-organizing (SO) systems have the capabilities to induce low-dimensional global behavior in a high-dimensional system, for instance in bird flocks or ant colonies. These systems tend to be resilient under changeable conditions or to malfunction of individual parts. In the case of agents' behavior, the self-organization relies on the embodied interaction with the environment and exploits physical effects. These properties make it natural to aim for the application of a self-organizing method to robot development. A difference between the prominent SO systems and a robot is that it consists of a single component instead of many largely independent components. In order to obtain a self-organized behavioral development, a guiding principle is required. One of the prominent principles is Homeokinesis – balancing activity and predictability. The maximization of the predictive information (PI) [1] is another such principle that was also cast into a online algorithm [3]. Although these methods lead to self-organization of behavior in a large variety of robots, the principle has an inherent disadvantage of constantly increasing the complexity of the behavior. As an effect the system aims for behaviors with maximal attractor dimension, see [5]. However, natural behavior such as walking, crawling, reaching, etc. are low-dimensional. Conceptually, it might be a question of timescales. If a succession of low-dimensional behaviors is generated then on a longer time-scale the predictive information can potentially also be maximized. Nevertheless, for direct behavior generation, it may ultimately not be suitable.

Recently the differential extrinsic plasticity (DEP) [2] learning rule was proposed, which originated from a simplification of the PI rule. It leads to the emergence of low-dimensional behaviors in different systems. Examples are a humanoid robot starting to crawling, turn a wheel or a hexapod robot to locomote [2]. Applied to a real tendon driven anthropomorphic arm the learning rule led to the SO development of bottle swinging, and shaking, table wiping and wheel turning, just due to the brain-body-environment coupling [4]. The DEP rule creates a dynamics with a multitude of attractors. Depending on the body and environmental conditions different attractor behaviors are reached. These can be switched by perturbations. However, a systematic understanding is still missing. In this contribution we provide a first analysis of DEP rule and propose a method to obtain a systematic sweep through the behavior landscape.

Differential Extrinsic Plasticity (DEP)

We consider a robotic system. Let $x_t \in \mathbb{R}^n$ be the vector of sensor values at time t. The motor commands $y_t \in \mathbb{R}^m$ are given by a one-layer neural network: $y_t = \tanh(\kappa \frac{C_t}{\|C_t\|+\lambda} x_t)$ with the normalized synaptic strength matrix C. $\lambda \ll 1$ is a regularization parameter. The synaptic strengths are adapted according to the DEP rule as

$$\tau \Delta C_t = \langle \tilde{\dot{y}}_t \dot{x}_t \rangle - C_t \,, \tag{1}$$

where $\tilde{\dot{y}}_t = F(\dot{x}_{t+1})$ with F being an adaptive inverse model, relating sensor values to actions.

It was shown in [2] that oscillatory behaviors are stationary solutions of the dynamics. There are two main parameters: κ regulates the norm of the acting synaptic matrix, thus controlling the amplitude of the oscillations and how strongly the neurons are in their saturation region; the second parameter τ controls the time scale of the synaptic dynamics. Intuitively it controls how many momentary correlation terms $\tilde{\dot{y}}_t \dot{x}_t$ are effectively contributing to the synaptic matrix C.

For a systematic understanding of the DEP rule we consider a short circuit setup. In this setup the robot is replaced by a "short circuit" such that the next sensor values are given directly by the motor actions, i. e. $x_{t+1} = y_t$. Thus, also the inverse model can by replaces by the unit mapping: $F(x) = \mathbb{I}x$.

Attractors and their Basins

In principle the dynamical system in the 2-dimensional short circuit setup ($n = 2$) is already at least 8 dimensional, because the state is given by x, its derivative $\dot{x}$ and C. In practice, for the robotic experiments, the initial condition

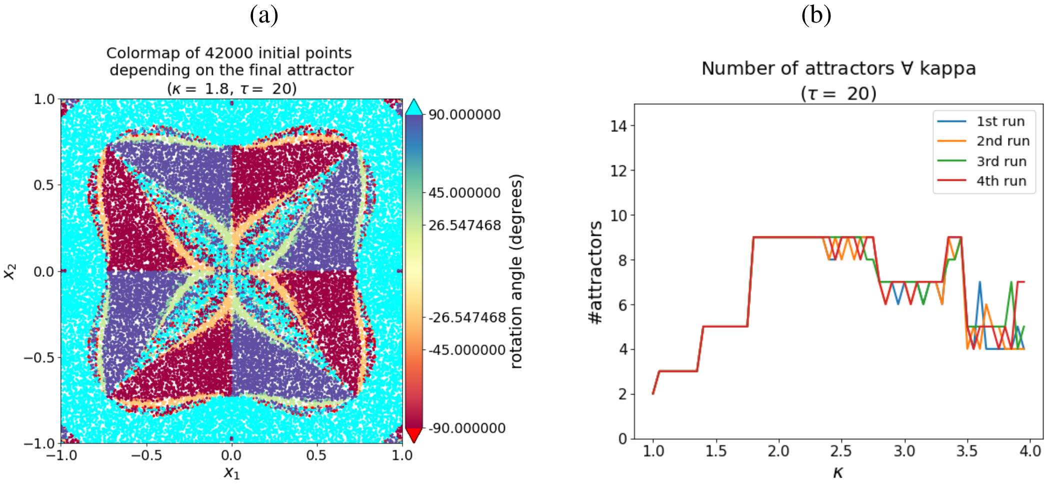

Figure 1: Number of Attractors and their basin of attraction. (a) Attractor behavior depending on starting position x for $\kappa = 1.8$ and $\tau = 20$. Note the regions with tiny basin of attractions. The color codes for the rotation angle between subsequent sensor states x. Red denotes $x = 0$ and cyan the period-2 orbits. (b) Number of different attractor behaviors depending on the parameter κ for $\tau = 20$ determined by a linkage clustering analysis.

was $C = \mathbf{0}$, such that all initial entries in the controller matrix (C) arise from the initial sensor configuration and environmental feedback included in $\dot{x}$ and $\tilde{y}$. To emulate this in the short circuit setting we consider the dynamics when starting from different x. During the first 3 steps there is not enough information to compute the derivative in Eq. (1), such that we follow the dynamics with $C = \mathbb{I}$. Afterwards, the usual update equations are evaluated.

The dynamics settles into stable attractors, either with a fixed C matrix, or with a quasiperiodic oscillation of C with a very small amplitude, which is treated as a fixed C matrix. The attractors are most decisively characterized by the dynamics of the sensor vector x. There is one fixed point $x = 0$, two period-2 oscillations, two period-4 oscillations and several quasiperiodic osciallations with slower frequency, see Fig. 1(a). The basis of attraction for each attractor are varying in size and there are regions with tiny basins. In this region the system is highly sensitive to the initial conditions. We studied the number of attractors depending on the parameter choice, an example is given in Fig. 1(b). Analogously, we performed this in higher dimensions. There is a wide range of values for κ with many attractors. In the robotic applications, the sensitivity to initial conditions allows the system to be sensitive to feedback from the physical system, such as faint footprints of a latent oscillations. The fact that the system has many attractors, even without any physical system attached, gives hints on why the rule can excite many different behaviors and that these can be easily switched by external influences.

Active switching of behavior – systematic sweep of attractors

The advantage of attractor behaviors is their stability. However, in view of a self-determined exploration of a robotic system, it would be great to actively explore all attractor behaviors in a systematic fashion. In this contribution, we will highlight our new approach to achieve this. Considering the DEP rule (1) as a differential equation we can add the derivative of a potential V that is constructed in a way to avoid all visited attractors in the space of C: $\tau\dot{C} = \langle\dot{\tilde{y}}\dot{x}\rangle - C + \frac{\partial}{\partial C}V(C, \mathcal{C})$, where $\mathcal{C}$ is the set of all attractors to be avoided. The potential is constructed by a sum of gaussians around each element of $\mathcal{C}$. In order to avoid the creation of spurious attractors, a particular merging of repellors is performed, such that the potential landscape is not excessively deformed. In preliminary experiments we find that the new method finds all attractor behaviors in a sequential fashion, as shown in Fig. 2. All original attractors (black dots) are visited one by one. The automatic repellor placement is indicated by the empty circles.

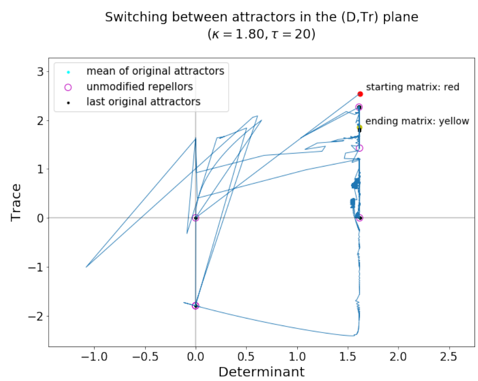

Figure 2: Active switching of attractors, here shown, for visualization purposes, in the determinant-trace plot of the C matrix, over time.

Conclusions

We have presented our recent advances in the analysis of the recently proposed DEP learning rule. Through the systematic study of a simplified system, we can find support for the observed behavior in the robotic experiments [2, 4]. Our newly proposed method can be used to systematically switch behavior. A practical evaluation in robotic experiences is left for future work.

Acknowledgement: We thank Ralf Der for stimulating discussions.

References

[1] N. Ay, N. Bertschinger, R. Der, F. Güttler, and E. Olbrich. Predictive information and explorative behavior of autonomous robots. *The European Physical Journal B*, 63(3):329–339, 2008.

[2] R. Der and G. Martius. Novel plasticity rule can explain the development of sensorimotor intelligence. *Proceedings of the National Academy of Sciences*, 112(45):E6224–E6232, 2015.

[3] G. Martius, R. Der, and N. Ay. Information driven self-organization of complex robotic behaviors. *PLoS ONE*, 8(5):e63400, 2013.

[4] G. Martius, R. Hostettler, A. Knoll, and R. Der. Compliant control for soft robots: emergent behavior of a tendon driven anthropomorphic arm. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 767–773, 2016.

[5] G. Martius and E. Olbrich. Quantifying emergent behavior of autonomous robots. *Entropy*, 17(10):7266, 2015.