Scaling of Guided Self-Organized Behavior

Georg Martius
Max Planck Institute for Mathematics in the Sciences,
Inselstr. 22, 04103 Leipzig, Germany
georg@mis.mpg.de

Introduction

Learning in autonomous agents implies an active involvement of the agent in the acquisition of new behavior. An effective exploration is essential for a successful learning in high-dimensional systems. Intrinsic motivation may the source of this exploration and lead to a new perspective for developmental robotics [3]. While a number of examples exist that impressively demonstrate the virtues of this view, it appears that non of them is able scale up to continuous domains with many degrees of freedom.

We will use here an approach to guided self-organization [5] in order to improve the sensitivity with respect to given learning signals. We will show examples where the learning time within this approach scales very nicely with the complexity of the problem.

Self-organizing behavior

Efficient behavioral exploration can be obtained by the homeokinetic principle [1, 2], a dynamical systems approach to robot control that establishes a self-tuned balance between sensitivity of actions to sensory inputs and predictability of the perceptual consequences of actions. The principle gives rise to a synaptic plasticity rule for artificial motor neurons such that coherent movements are generated that are suitable to explore the behavioral manifold [4]. Simultaneously, an internal representation of the robots dynamics is learned by a neural network in the robot. From this control principle emerge behaviors, that follow no specific goals, but exploit the embodiment [7] to produce suitable sensorimotor couplings.

Guided self-organization

How can we guide the learning dynamics such that a given goal is realized by the self-organizing process? Combining the self-organization of behavior and a task-related objective is a very promising approach, called guided self-organization (GSO) [6, 8]. While the homeokinetic principle explores behaviors that correspond to controllable sensorimotor loops, the additional requirements specify a subset of those. Pure self-organization may correspond to an early phase of behavioral development which, however, wanders through the set of sensitive sensorimotor couplings that allow for goal-dependent behaviors without reducing the self-induced activity of the robot.

Guidance by cross-motor teaching A first and simple way of guidance is by direct teaching signals for either motor values or sensor values [5]. In the present contributions we report on an approach that enforces structural properties, such as symmetries, of the desired behaviors by internal teaching signals. The idea is that the target value for one motor neuron is provided by the value of another motor neuron, thus we call this mechanism *cross-motor teaching*.

As a first example we want to influence the controller to prefer a mirror-symmetry in the motor patterns. This can be achieved by using the motor value of one motor as the teaching signal for another motor and vice versa. This self-supervised teaching induced soft constraints which reduce the effective dimension of the sensorimotor dynamics and thus guide the self-organization along a sub-space of the original control problem.

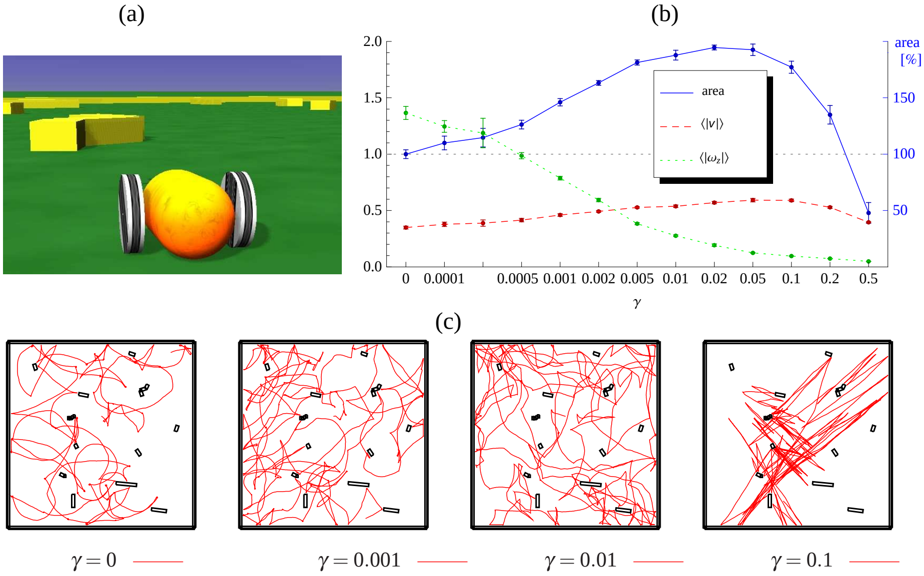

Figure 1: Behavior of a TwoWheeled robot that it guided to move preferably straight. The robot is placed in an environment cluttered with obstacles. **(a)** Screen shot from the simulations. **(b)** Mean and standard deviation (of 5 runs each 20 min) of the area coverage ('area'), the average velocity $\langle|\boldsymbol{v}|\rangle$, and the average angular velocity $\langle|\omega_z|\rangle$ for different values of the guidance factor γ. Area coverage (box counting method) is given relative to the case without guidance, i.e. 100% for $\gamma=0$ (right axis). At increasing γ the robot is moving straighter and its trajectory covers more area, until at large γ the teaching dominates the behavior of the robot. **(c)** Example trajectories for different guidance factors.

Let us consider a two-wheeled robot and suppose the robot should move mostly straight, not get stuck at obstacles or in corners and cover substantial parts of its environment. We will see that all this can be achieved by a simple guidance of the homeokinetic controller where both motor neurons are mutually teaching each other, see Fig. 1. Depending on the strength of guidance the robot shows a distinct decrease in mean turning velocity and a higher area coverage until the guidance dominates the behavior. The robot is still performing turns and drives both backwards and forwards and does not get stuck at the walls, as seen in the trajectory in Fig. 1(c),

Let us now consider a more complicated high-dimensional example using the Armband robot, see Fig. 2(a). The robot has 13 degrees of freedom and only 13 joint angle sensors. The robot is to locomote in a particular direction. This can be achieved by a cyclic cross-motor teaching setup, where motor neurons driving joints at opposite sides of the robots teach each other in a directional way. For a suitable strength of the guidance, the robot develops after about 1 min a locomotion behavior from scratch, which involves a cooperation of all joints. The velocity of the robot depends on the guidance strength. For overly strong guidance the behavior gets worse because the self-organizing process is disturbed too much. The direction of locomotion can also be inverted by changing the motor connection setup, cf. Fig. 2(b).

Scaling properties

The locomotion of the robot is essentially influenced by the number of cross-motor connections. In a series of simulations a number $0 \leq l \leq m$ equally spaced cross-motor connections (Fig. 2(a)) are used. With increasing l the robot start to locomote earlier. Full performance is reached already if 8 out of the 13 connections are used, see Fig. 3(a).

In order to study the scaling properties of the learning algorithm we varied the number of segments m of the robot and thus the dimensionality of the control problem. The results are astonishing, see Fig. 3(b): The behavior is learned with the same speed also for large number (40) of segments. There is no scaling

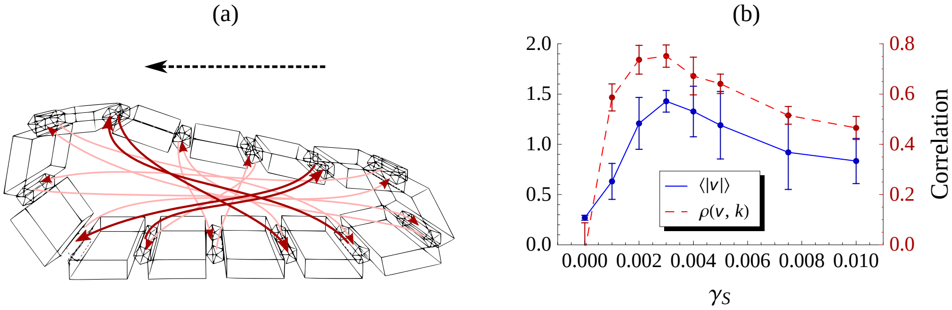

Figure 2: Locomotion of a high-dimensional robot with cross-motor teaching. **(a)** Schematic view of the robot. The prismatic structures represent the actuated joints. The arrows indicate the cross-motor teaching connections (the thick arrow show the case of $l = 4$, for $l = 13$ all are equal in strength). **(b)** Mean and standard deviation of the average absolute velocity $\langle|v|\rangle$ and the correlation $\rho(v,k)$ of the velocity with the configuration of the coupling $k \in \{\text{forward}, \text{backward}\}$ for 5 runs each of 30 min for different guidance strengths γ_S. The coupling configuration was changed every 5 minutes. The correlation does not reach unity because new behaviors are not perfectly learned and nearby behaviors are constantly explored.

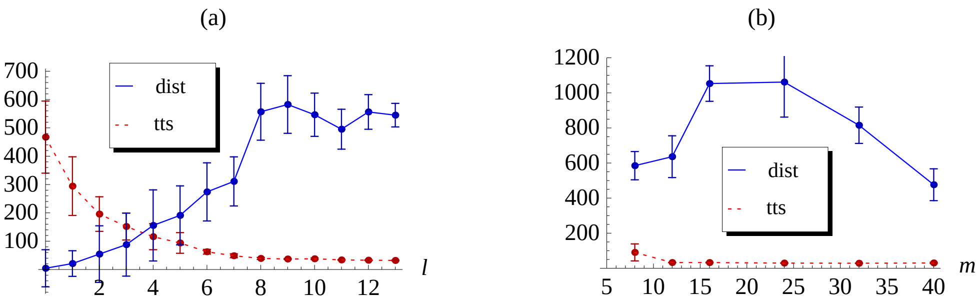

Figure 3: Scaling of learning time and performance for different robot complexity. The plots show mean and standard deviation of the distance traveled by the robot (‘dist’ in units of 1 segment size) and of the time-to-start (‘tts’ in seconds) of 20 runs à 10 min ($\gamma = 0.003$). **(a)** Performance as a function of the number of cross-motor connections l (equally spaced around a robot with $m = 13$ joints). **(b)** Performance for different numbers of segments m (DoF) with full cross-motor connectivity ($l = m$).

problem here for the following reason. In the closed loop with an approximate feedback strength (self-regulated by the homeokinetic controller) the robot needs only very little influence to roll. The length of the robot can even help because other behavioral modes (e. g. wobbling) are damped increasingly due to gravitational forces. For the same reason, small robots are slower than medium ones. Large robots are again slower because the available forces at the joints become too weak. The experiment illustrates that specific behaviors can be achieved in a high-dimensional robot by using cross-motor teachings. Cross-motor connections can break the symmetry between the two directions of motion such that a locomotory behavior is produces quickly. When the connections are switched later during runtime, the behavior of the robot changes reliably.

Discussion

Constraining the behavioral self-organization by fixed objectives is seen to provide reasonable solutions if the objective function specifies a lower dimensional subspace. This subspace characterizes all behaviors that are compatible with the objective and can be generated by the homeokinetic controller of the robot. Guidance by teaching signals and cross-motor teaching are implemented using an explicit gradient decent, which makes the approach very fast. The exploitation of the physical constraints through the self-model and the sensitive control regime lead to a successful learning also in high-dimensional systems. Importantly there is no scaling problem in the considered cases, such that a system of 40 degrees of freedom learned a locomotion behavior within minutes.

In the context of animal development the influence of learning by self-organization and by reward mechanisms may vary. Although in the early stages the pure self-organization of sensorimotor loops can be expected to follow mainly intrinsic principles, later stages will see a combination of different learning mechanisms as described here in an exemplary case. We find that the maintenance of criticality that is essential in the homeokinetic approach is not abandoned with goal-oriented learning as rather weak effects of the objective are most efficient. It can, furthermore, be predicted that an early exploratory phase which is not subject to directed learning increases the efficiency with which later the objective is met.

Acknowledgements: This work was supported by the BMBF grants 01GQ0432.

References

[1] R. Der. Self-organized acquisition of situated behaviors. *Theory in Biosciences*, 120:179–187, 2001.

[2] R. Der and R. Liebscher. True autonomy from self-organized adaptivity. In *Proc. of EPSRC/BBSRC Intl. Workshop on Biologically Inspired Robotics*, HP Labs Bristol, 2002.

[3] M. Lopez and P.-Y. Oudeyer. Active learning and intrinsically motivated exploration in robots: Advances and challenges. *IEEE Transactions on Autonomous Mental Development*, 2(2):65–69, 2010.

[4] G. Martius, K. Fiedler, and J. M. Herrmann. Structure from Behavior in Autonomous Agents. In *Proc. IEEE Intl. Conf. Intelligent Robots and Systems (IROS 2008)*, pages 858 – 862, 2008.

[5] G. Martius and J. M. Herrmann. Taming the beast: Guided self-organization of behavior in autonomous robots. In S. Doncieux, B. Girard, A. Guillot, J. Hallam, J.-A. Meyer, and J.-B. Mouret, editors, *From Animals to Animats 11*, volume 6226 of *LNCS*, pages 50–61. Springer, 2010.

[6] G. Martius, J. M. Herrmann, and R. Der. Guided self-organisation for autonomous robot development. In F. Almeida e Costa, L. Rocha, E. Costa, I. Harvey, and A. Coutinho, editors, *Proc. Advances in Artificial Life, 9th European Conf. (ECAL 2007)*, volume 4648 of *LNCS*, pages 766–775. Springer, 2007.

[7] R. Pfeifer and J. C. Bongard. *How the Body Shapes the Way We Think: A New View of Intelligence*. MIT Press, Cambridge, MA, November 2006.

[8] M. Prokopenko. Guided self-organization. *HFSP Journal*, 3(5):287–289, 2009.