

Embodied Evolution and Learning: The Neglected Timing of Maturation

Steffen Wischmann, Kristin Stamm, and Florentin Wörgötter

Bernstein Center for Computational Neuroscience,
University of Göttingen, Bunsenstr. 10, 37073 Göttingen, Germany,
{[steffen|kstamm|worgott](mailto:steffen|kstamm|worgott}@bccn-goettingen.de)}@bccn-goettingen.de

Abstract. One advantage of the asynchronous and distributed character of embodied evolution is that it can be executed on real robots without external supervision. Further, evolutionary progress can be measured in real time instead of in generation based evaluation cycles. By combining embodied evolution with lifetime learning, we investigated a largely neglected aspect with respect to the common assumption that learning can guide evolution, the influence of maturation time during which an individual can develop its behavioral skills. Even though we found only minor differences between the evolution with and without learning, our results, derived from competitive evolution in predator-prey systems, demonstrate that the right timing of maturation is crucial for the progress of evolutionary success. Our findings imply that the time of maturation has to be considered more seriously as an important factor to build up empirical evidence for the hypothesis that learning facilitates evolution.

1 Introduction

Evolutionary Robotics (ER) has become a well established method for investigating the development of adaptive behavior in situated and embodied agents (e.g., [1,2]). However, demonstrating truly open-ended evolution of complex behavior is still a great challenge [3]. A regularly raised argument for that is the strong simplification of ER experiments with respect to natural evolution and the complexity of biological organisms (for a discussion see, for instance, [4]). Although this criticism has to be taken seriously, it seems rather unpractical to begin with.

Instead of trying to increase the complexity of individual agents at any cost, we propose a more pragmatic approach. The aim of this paper is to present a methodology which unifies several aspects of ER research, which previously have been considered rather separately: the interplay between ontogenetic (learning) and phylogenetic (evolutionary) development [5], the evolution of morphology and control [6], the use of distributed and asynchronous evolutionary algorithms [7], the role of competitive co-evolution [8], and the reduction of fitness function constraints to achieve open-ended evolutionary processes [4]. We argue, unifying those aspects gives promising ground for the development of more biologically

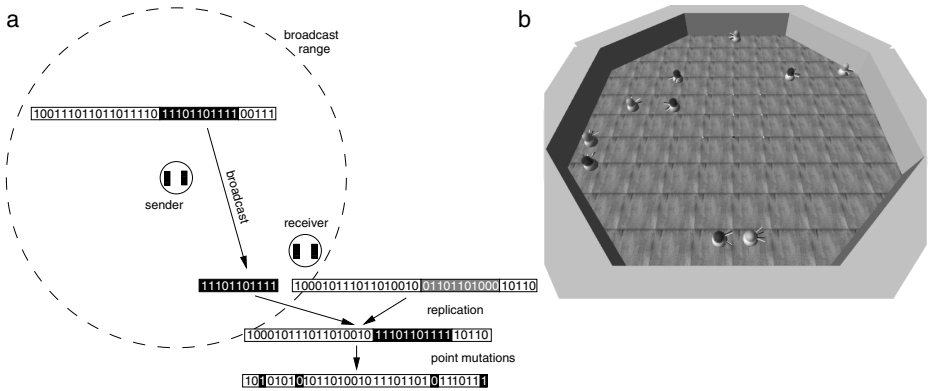


Fig. 1. a: Principle mechanisms of the embodied evolutionary algorithm (see text for details). b: Physical simulation of a predator-prey scenario (two populations can be distinguished by the color on top of the robots).

plausible ER experiments, which may allow a successful incremental increase of behavioral complexity.

Besides the description of such a unified methodology, we present first experiments involving competing populations of predators and preys. Our results challenge the previously hold assumption that learning facilitates evolution [9,5,10]. This assumption is mainly based on experiments where, as we believe, an important aspect of lifetime adaptation has been largely neglected—the time of maturation during which an individual can develop its behavioral skills. Whether or not learning facilitates evolution crucially depends on this factor, as we can show within our experimental framework. If the timing is wrong, learning can even be counterproductive and individuals are better off with evolutionary development only. However, with the right timing, learning can slightly improve evolutionary success. We argue further, to build empirical evidence for the hypothesis that learning can guide evolution in the context of ER, we have to abandon classical generation based evolution and should instead consider an asynchronous and distributed approach as pursued in this paper.

2 Methods

In the following we consider a predator-prey scenario involving two competing populations. In remembrance of the pioneering work of Werner and Dyer [11] we call prey individuals sheep and predators wolves. Sheep can gather energy by moving around without getting caught by wolves (an analogy to grazing), whereas wolves gain energy by catching sheep. Here, only the sheep population evolves whereas wolves can only perform lifetime adaptation. However, the applied evolutionary algorithm (EA) can be easily used to investigate co-evolution

of both populations which was omitted here to provide a better analysis of the fitness development by eliminating the Red Queen effect [12], that is, an evolutionary arms race as investigated in [13].

Embodied Evolution. The EA we implemented here is a slightly modified version of the probabilistic gene transfer algorithm (PGTA) introduced by Watson et al. [7]. It is an asynchronous and distributed *embodied* EA, which can be executed directly on real robots without the need of external supervision. Robots of a population evolve by exchanging gene strings of their genome dependent on their current performance. Each robot executes the following algorithm:

```

DEFINE embodied_evolve
  initialize_genes[]
  energy = initial_energy
  REPEAT FOREVER
    IF (maturation_exceeded?){
      IF (excited?)
        send(genes[random(num_genes)])
      IF (receptive? and received?){
        genes[indexof(received)] = valof(received)
        mutate(genes[])
        begin_maturation()
        energy = maturation_energy}}
    do_task_specific_behavior
    energy = limit(energy + reward - penalty)
  ENDREPEAT
ENDDDEFINE

```

At the beginning of evolution, the genome of a robot is initialized randomly and a starting energy is set (this can be either the actual battery current or a virtual energy value). In every time step, depending on the current energy, it is probabilistically determined if a robot sends a sequential part of its genome (*excited*). The gene string can be received by all individuals within a predefined broadcast range (see Fig. 1a). A robot accepts received gene strings with probability $(1 - \text{energy})$ and overwrites part of its genome with the new string. The new genome is mutated with a predefined mutation rate. With gene acceptance the maturation period begins and the energy level is reset. Maturation period is a developmental phase where individuals can neither broadcast nor receive gene strings. Note, this period does not determine the lifetime of an individual, it just assures a minimum time of development. The actual lifetime depends on the performance of an individual, the better it is the longer the individual lives. At the end of each time step the energy is updated. If the robot broadcasts a gene string, energy is decreased with a penalty (in analogy to reproduction costs). If the robot successfully accomplishes its task, it gains an energy reward. The energy of a robot is limited by a minimum and maximum value.

The major difference to the PGTA proposed in [7] is the maturation period which ensures that new genomes are able to prove their performance potential

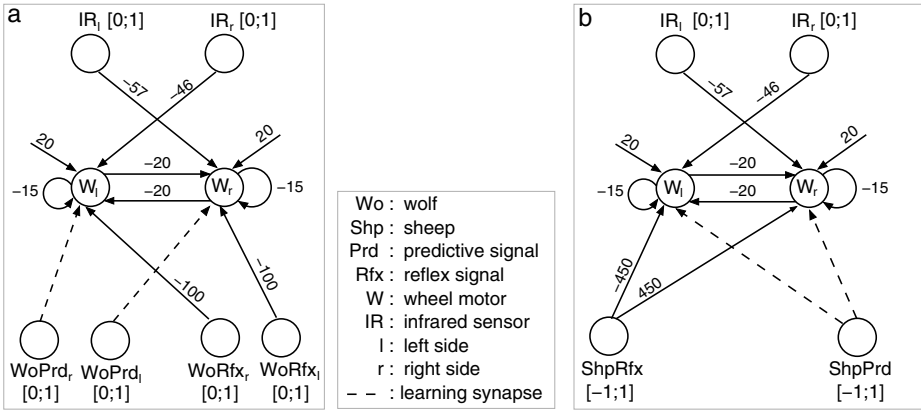


Fig. 2. Neural networks controlling sheep (a) and wolves (b). Square brackets enclose the range of neuronal output (W_l and W_r are within $[-1, 1]$).

before they may get overwritten by another received gene string. Further, mutation is applied to the whole receiver's genome instead of to the transferred gene string only. This allows more variation in new genotypes.

Learning. During lifetime wolves and sheep are able to adapt their behavior—either catching sheep or escaping from wolves. As adaptation mechanism we applied isotropic-sequence-order learning using input correlations only (ICO) described in [14]. This approach is based on temporal correlations between sensor signals as they appear in nature, too. For instance, a child learns to react to heat perception caused by a hotplate which precedes the pain when touching it. In our example, seeing a wolf precedes getting caught. Sheep react to wolf attacks by trying to move away. This is an inbuilt reflex. During lifetime, they are able to learn to prevent these reflexes by using predictive sensory signals—e.g. sensing of approaching wolves. ICO learning realizes correlations between a reflexive and a preventive behavior. Preventive synaptic weights in the neural control are learned by the following learning rule:

$$\frac{d}{dt}p_i = \mu u_i \frac{du_r}{dt}, i \in [0, N], \quad (1)$$

where a weight change of a synapse p_i in time step t depends on the learning rate μ , the input signal u_i transmitted by the learned synapse, and the difference in a reflex input signal u_r (for details see [14]).

Fig. 2 presents the neural network (NN) control including ICO learning of sheep and wolves. Both NNs include a robust obstacle avoidance functionality realized by infrared (IR_l, IR_r) sensor signals influencing the speed of the left and right wheel (W_l, W_r). The IR signals detect obstacles on the left and right frontal side of the robot. An obstacle detected on the robot's left side causes

a right turn and vice versa. The cyclic interconnections of the wheel neurons cause a hysteresis effect to enable handling of acute angles (for details see [15]). Additionally, sheep have a predefined wolf aversive reflex, implemented by reflex sensor signals ($WoRfx_l, WoRfx_r$), that works similar to obstacle avoidance. Reflex sensors are only triggered if a sheep collides with a wolf. To prevent these collisions sheep have also predictive signals ($WoPrd_l, WoPrd_r$) which recognize wolves within a larger range. At the beginning, the synaptic connections to the wheel neurons have zero weight, but can be learned by using ICO. Wolves instead have an inbuilt pursuit reflex. The reflex signal $ShpRfx$ transmits zero if the sheep is in front of the wolf, but negative or positive values, if the sheep is on the left or right side, respectively. The wolf turns towards the sheep until the sheep is in front of it. But, $ShpRfx$ transmits only signals if the sheep is within a small collision range. Wolves have also a prediction sensor $ShpPrd$ that functions in the same way as $ShpRfx$, but senses sheep within a larger range. $ShpPrd$ synaptic connections can be learned with ICO.

Competitive Evolution. For the implementation of embodied evolution we have selected evolvable parameters of sheep from different domains: neural control, embodiment, and lifetime adaptation. We evolved maximum wheel motor speed, position of the wolf detection sensors (similar to the sensor position evolution described in [16]), and the ICO learning rate μ . Each parameter is restricted by a value range and represented by a 10 bit string in the binary genome (see [17] for details).

Instead of defining high level targets, the fitness function simply corresponds to the energy level of a robot, that is, it describes a rather general survival criteria, as it is considered to be essential for open-ended and creative evolution [3,4]. Each sheep gains energy whenever there is no wolf within their collision range, whereas a wolf gains energy while a sheep is within its collision range (i.e., the energy of a sheep is transferred to the wolf). Additional energy loss is realized dependent on the broadcast behavior. Anytime an individual broadcasts a gene string, energy is reduced (in analogy to reproduction costs).

For the ease of analysis we implemented the robots and the environment with the physical simulation library *ODE* (see Fig. 1b) based on the specifications of the KheperaTM robot¹. For the following experiments each population of sheep and wolves consists of three individuals where sheep can evolve and wolves can adapt their behavior by learning only.

3 Results

3.1 Maturation Time

First of all we wanted to investigate the influence of the time an individual has to develop until it becomes exposed to the selection pressure (i.e. its maturation

¹ The limited space of this paper prohibits detailing all parameters, but a precise description of the simulation can be found in [17].

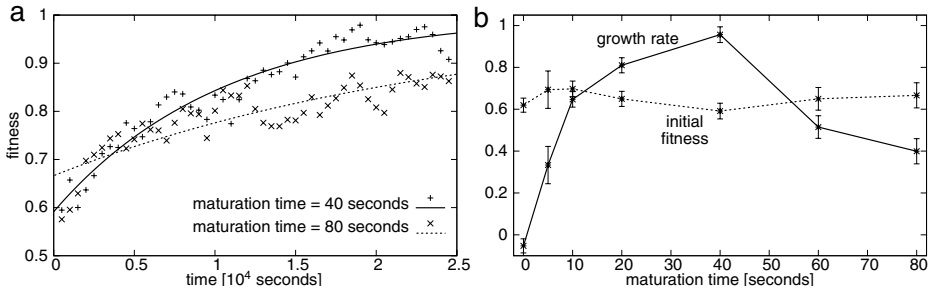


Fig. 3. a: Data points indicate fitness development for two distinct maturation times (see text for calculation method). Data is approximated by an exponential function (see text for details). b: Growth rate and initial fitness depending on maturation time (error bars indicate the mean error of the approximated exponential function).

period). For each of several distinct maturation times (MTs) we conducted 25 evolutionary runs with randomly varied initial conditions (i.e., randomly initialized genomes and random starting positions). Each run lasted 2.5 million time steps (time step = 0.01 sec) which corresponds to approximately seven hours of real time, if we would execute it on real robots. This is a reasonable time scale for embodied evolution experiments (see also [7]). During these 25 runs we recorded the energy levels of all robots in each time step. We then took the mean of all robots and all runs. To smoothen sudden jumps in the data which can occur if a sheep gets caught or receives a new genome, we averaged again over a time window of 500 seconds.

Fig. 3a shows the resulting data for two different MTs. To get a clearer picture of the fitness development during evolution, we approximated the data by an exponential function of type $f(x) = 1 - ae^{-bx}$, where $(1 - a)$ indicates the initial fitness (i.e., after the first 500 sec.) and b the growth rate over time (an unconstrained nonlinear optimization method of the MatlabTM optimization toolbox was used). In Fig. 3a we see that individuals with the shorter MT (i.e., 40 sec.) do not only increase their fitness much faster than individuals with a longer MT (i.e., 80 sec.), but they also reach nearly maximal fitness already after 18,000 seconds.

Fig. 3b illustrates the approximated growth rate and initial fitness for different MTs. The initial fitness is nearly constant. The growth rate, however, differs significantly. We find a maximum at 40 seconds, and declining growth rates for smaller and larger MTs. Explaining smaller growth rate at smaller MT is rather straight forward. Individuals do not have time enough to adapt their learning synapses before they are exposed to the selection pressure. The evolutionary progress is rather slow because this holds for all evolving individuals and even though the frequency of gene exchange is rather high, the fitness difference between sender and receiver is not a reliable indicator for a better or worse performance. With short MT, some individuals may have encountered wolves a

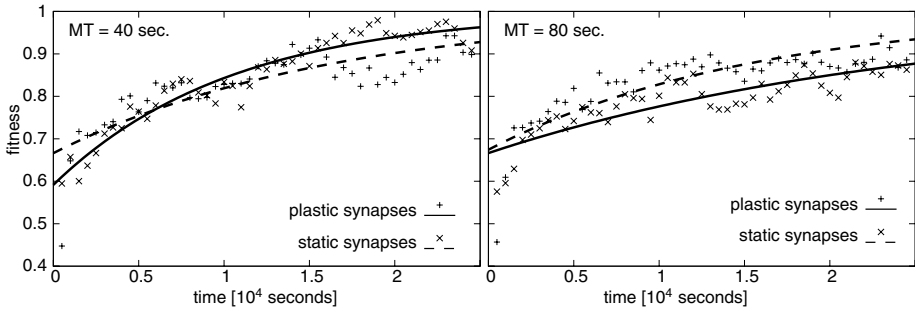


Fig. 4. Comparing fitness development during the evolution of individuals with plastic synapses and static synapses for different maturation times (MTs).

few times, but other may have never encountered them and, therefore, have a higher fitness. Thus, genes may be transmitted which are only successful because they never had to prove their performance (note that sheep have to encounter wolves to develop their escaping skill). This effect cancels out over time, but the evolutionary progress is slowed down. In contrast, for long MTs the gene exchange frequency is too slow. Even though individuals now have enough time to develop, evolving parameters (e.g., learning rate or morphological properties) which improve the performance spread only slowly throughout the population.

3.2 Learning and Evolution

To figure out how learning influences the course of evolution we conducted the same experiments as described before for neural networks with static synapses. Basically, nothing changed except that the weights of the synapses, which were learned during the lifetime of an individual (cf. Fig. 2a), were now encoded in the genome as well. Note, this was applied for sheep only, the setup for the wolf population remained the same, that is, wolves could learn, but were not evolved.

We chose two MTs, the one with the largest growth rate (40 sec.) and one with a small growth rate, where we preferred 80 to 5 seconds (cf. Fig. 3a) because the former provides more reliable fitness estimates than the latter (for the reason that sheep encounter more wolves during their evaluation). Fig. 4 compares the fitness development during the evolution with the two MTs to our previous results from the evolution with synaptic plasticity (same measurements were applied as described before). Considering only static synapses, we find no influence of MT. Compared to the case of plastic synapses, for an MT of 80 seconds the fitness development for evolution without learning is even slightly better than for evolution with learning. If, however, MT is shorter (i.e., 40 sec.), evolution with learning performs slightly better than without learning.

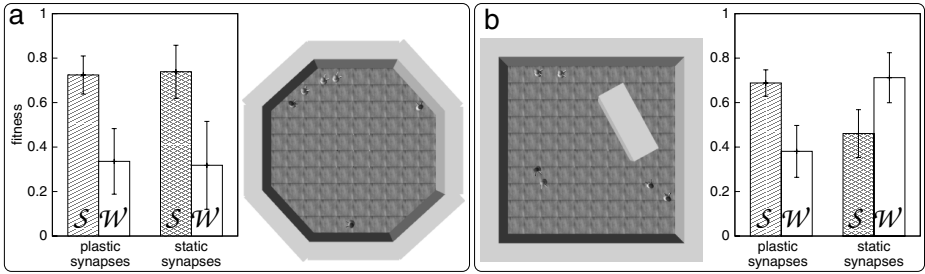


Fig. 5. Comparing performance of sheep populations (S) evolved either with plastic or static synapses in the environment with which they were evolved (a) and a novel environment (b). Performance of wolf populations (W) is drawn as a reference. Shown is the average and standard deviation of 25 randomly initialized runs.

3.3 Adaptability to Novel Environmental Conditions

In the previous section we have seen that learning may slightly guide evolution, but the effect within our experimental framework is not really significant. The good news is that both methods take the same computational time due to the asynchronous and distributed evolution approach. Thus, we may conclude that in those basic tasks it does not matter whether we evolve with or without synaptic plasticity (presupposing the right timing of maturation). However, the commonly assumed advantage of learning is that individuals may be able to adapt to significantly novel environmental conditions with which they were not confronted during their evolutionary development, as for instance shown in [10,18].

To test whether or not this assumption holds for our approach as well, we chose one representative evolutionary run from the evolution with and without learning with an MT of 40 seconds (cf. Fig. 4). From these two runs we took the neural networks of the sheep population at the end of evolution (i.e., after 25,000 sec.) and tested them first in the environment with which they were evolved for 400 seconds (i.e., ten times the original MT) while disabling genome transmissions. Fig. 5a illustrates the performance of a sheep population with either plastic or static synapses. As could be expected from Fig. 4, both solutions perform similar well. However, confronted with a significantly different environment, the population with plastic synapses performs significantly better (see Fig. 5b). This can be accounted to the adaptation mechanism which individuals with static synapses lack. And adaptation is required in this environment because compared to the original environment (cf. Fig. 5a), it contains sharp corners and narrow passages. Thus, sheep are much better off with reacting earlier to approaching wolves to avoid being captured in corners. Optimal reaction time was not required in the original environment. There, the most important parameter was the speed of sheep which was always maximized by evolution. Thus, even with a moderate reaction time, sheep could escape approaching wolves for most

of the time. Confronted with the new environment learning individuals adjusted their sensory sensitivity with respect to approaching wolves by a weight increase of the according synapses (see Fig. 2a). Therefore, if environmental conditions change, individuals with synaptic plasticity are able to react to those changes and maintain their original performance, whereas non-learning individuals are not able to cope with this change.

4 Conclusions

The very nature of embodied evolution forces us to abandon single individual evolution. Even under basic environmental conditions as considered here, robotic agents already have to deal with a dynamically changing environment. Even further, right from the beginning they have to develop robust behaviors because it is not guaranteed that other individuals of their own species, they may encounter, act reasonable. Due to the asynchronous and distributed character of embodied evolution, conspecifics can show quite different behaviors at the very same time. For instance, some individuals may try to avoid their conspecifics while others got some mutations which may lead them directly crash into their conspecifics.

The most intriguing aspect of a methodology as proposed here is the ability to incrementally increase behavioral complexity by increasing environmental factors. Here, we considered only one evolving population, but we also tested co-evolution of both populations and discovered an evolutionary arms race as already described in [13] for predator-prey systems with one individual in each population only. To do so, no change of fitness functions is required, robots ‘solely’ have to maintain their energy. For future research it would be interesting to consider more than two interacting populations by building up small ecosystems of intertwined species, each influencing the evolution of the other. Thus, we argue that our unified framework is suitable to approach more biologically plausible scenarios in ER striving towards open-ended and creative scenarios as proposed in [4,3].

Besides these general methodological facets, we investigated an aspect largely neglected in previous studies concerning the interplay of evolution and learning [9,5,10]. We demonstrated that whether or not learning may facilitate evolution depends on the right timing of individual maturation. Further, we showed that the choice of maturation time significantly influences the rate of progress during evolution. Even though our experimental results can not offer a general answer to the question how long learning individuals should be able to develop, they clearly show that we have to consider this aspect sincerely to build up empirical evidence for the commonly hold assumption that learning guides evolution. And the use of asynchronous and distributed embodied evolution should be preferred over the classical generation based evolution; not only because it can be executed directly on real robots, but also because it allows a direct comparison of evolution with and without learning measured on real time progress.

Acknowledgments. This work was supported by the PACO-PLUS project, funded by the European Commission under grant IST-FP6-027657.

References

1. Nolfi, S., Floreano, D.: *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press, Cambridge (2000)
2. Harvey, I., Di Paolo, E.A., Wood, R., Quinn, M., Tuci, E.: Evolutionary robotics: A new scientific tool for studying cognition. *Artificial Life* 11, 79–98 (2005)
3. Taylor, T.: Creativity in evolution: Individuals, interactions and environments. In: *Creative Evolutionary Systems*, pp. 79–108. Morgan Kaufmann, San Francisco (2001)
4. Bianco, R., Nolfi, S.: Toward open-ended evolutionary robotics: Evolving elementary robotic units able to self-assemble and self-reproduce. *Connection Science* 16, 227–248 (2004)
5. Nolfi, S., Floreano, D.: Learning and evolution. *Autonomous Robots* 7, 89–113 (1999)
6. Taylor, T., Massey, C.: Recent developments in the evolution of morphologies and controllers for physically simulated creatures. *Artificial Life* 7, 77–87 (2001)
7. Watson, R.A., Ficici, S.G., Pollack, J.B.: Embodied evolution: Distributing an evolutionary algorithm in a population of robots. *Robotics and Autonomous Systems* 39, 1–18 (2002)
8. Nolfi, S., Floreano, D.: Co-evolving predator and prey robots: Do ‘arm races’ arise in artificial evolution? *Artificial Life* 4, 311–335 (1998)
9. Hinton, G.E., Nowlan, S.J.: How learning guides evolution. *Complex System* 1, 495–502 (1987)
10. Urzelai, J., Floreano, D.: Evolution of adaptive synapses: Robots with fast adaptive behavior in new environments. *Evolutionary Computation* 9, 495–524 (2001)
11. Werner, G., Dyer, D.: Evolution of communication in artificial organisms. In: *Proc. Artificial Life II*, Addison-Wesley, Reading (1991)
12. van Valen, L.: A new evolutionary law. *Evolutionary Theory* 1, 1–30 (1973)
13. Floreano, D., Nolfi, S.: God save the red queen! Competition in co-evolutionary robotics. In: *Genetic Programming 1997: Proc. of the Second Annual Conference*, pp. 398–406 (1997)
14. Porr, B., Wörgötter, F.: Strongly improved stability and faster convergence of temporal sequence learning by utilising input correlations only. *Neural Computation* 18, 1380–1412 (2006)
15. Hülse, M., Wischmann, S., Pasemann, F.: Structure and function of evolved neuro-controllers for autonomous robots. *Connection Science* 16, 249–266 (2004)
16. Buason, G., Bergfeldt, N., Ziemke, T.: Brains, bodies, and beyond: Competitive co-evolution of robot controllers, morphologies and environments. *Genetic Programming and Evolvable Machines* 6, 25–51 (2005)
17. Stamm, K.: Individual learning and the dynamics in predator-prey populations. Master’s thesis, University of Göttingen (2007)
18. Di Paolo, E.A.: Organismically-inspired robotics: Homeostatic adaptation and teleology beyond the closed sensorimotor loop. In: *Dynamical Systems Approach to Embodiment and Sociality*, pp. 19–42. Advanced Knowledge International, Magill (2003)