Chapter 5
Evolving Monolithic Robot Controllers through

Joshua E. Auerbach and Josh C. Bongard

Abstract. Evolutionary robotics has been shown to be an effective technique for generating robot behaviors that are difficult to derive analytically from the robot's mechanics and task environment. Moreover, augmenting evolutionary algorithms with environmental scaffolding via an incremental shaping method makes it possible to evolve controllers for complex tasks that would otherwise be infeasible. In this paper we present a summary of two recent publications in the evolutionary robotics literature demonstrating how these methods can be used to evolve robot controllers for non-trivial tasks, what the obstacles are in evolving controllers in this way, and present a novel research question that can be investigated under this framework.

5.1 Introduction

What gives rise to intelligent behavior in natural and artificial agents? If you ask proponents of embodied artificial intelligence they will argue that such intelligent behavior arises out of the coupled dynamics between an agent's body, brain and environment [1, 6, 9, 17]. An extension of this idea is that the complexity of an agents's controller and morphology must match the complexity of the task or tasks that it is required to perform. However, when extending this idea to more complex agents in more complex environments it is not clear how to distribute responsibility for different behaviors across the agents's controller and morphology. Some have argued [8, 10] that controllers should be organized in a modular fashion such that different control components are responsible for different behaviors, but is this modularity

Joshua E. Auerbach
Morphology, Evolution and Cognition Laboratory, Department of Computer Science, University of Vermont, Burlington, VT 05405, USA
e-mail: joshua.auerbach@uvm.edu

Josh C. Bongard
Morphology, Evolution and Cognition Laboratory, Department of Computer Science, University of Vermont, Burlington, VT 05405, USA
e-mail: josh.bongard@uvm.edu

S. Doncieux et al. (Eds.): New Horizons in Evolutionary Robotics, SCI 341, pp. 55–65.
springerlink.com

necessary? Recent work by our group and others has demonstrated that in fact, no, structural modularity is not always necessary [2, 3, 7, 14]. An example of how a monolithic (non-modular) controller can be evolved to enable a virtual autonomous robot to perform a non-trivial sequence of behaviors will be presented in the next section.

Besides modularity in the design of an agent's controller, roboticists often implicitly design their robots to have morphological modularity as well: different parts of the robot's body are responsible for different behaviors. For example, wheels or legs may allow for movement while a separate gripper module allows for object manipulation. But, what if this assumption is relaxed? In another recent publication [2], we demonstrated how a robot could be trained to locomote to and manipulate an object while the assumption of specialization of different body parts is relaxed: the robot had a segmented body plan in which the front segment was able to participate in locomotion and object manipulation, or it might have specialized such that it only participated in object manipulation. In this way, selection pressure dictated the presence and degree of specialization of the robot's morphology rather than enforcing such specialization *a priori*. Section 5.3 summarizes this work and discusses some of the insight gained from studying the variability observed in the degree of specialization of evolved controllers across different experimental regimes.

5.2 Learning Multiple Behaviors with a Monolithic Controller

Evolutionary robotics [12, 16] has been shown to be an effective technique for generating robot behaviors that are difficult to derive analytically from the robot's mechanics and task environment. In particular, such techniques are useful for realizing dynamic behaviors (eg. [13, 18]) in which individual motor commands combine in a nonlinear fashion to produce behavior, thereby making analytical derivations of optimal controllers infeasible. However, evolutionary algorithms alone are usually insufficient for evolving controllers capable of multiple dynamic behaviors. One method of augmenting evolutionary algorithms to achieve such controllers is incremental shaping ([11],[19] and [20]): the gradual complexification of an agent's task environment, also known in the developmental psychology literature as scaffolding [21], in order to first train controllers capable of performing a simplified version of a given task and then over time increase the task difficulty.

In a recent publication [3] we showed how using an incremental shaping technique makes it possible to train a virtual autonomous robot to overcome three learning milestones: object manipulation, dynamic forward legged locomotion toward an object, and directed legged locomotion toward an object, all using a single monolithic controller. Moreover, that work demonstrated the necessity of choosing an appropriate shaping trajectory or scaffolding schedule opening up several questions about how to choose such a schedule.

Specifically, two virtual quadruped robots (see Fig. 5.1) simulated in a physically realistic simulation engine[1] were experimented with. Both robots had a desired task

[1] Open Dynamics Engine: www.ode.org

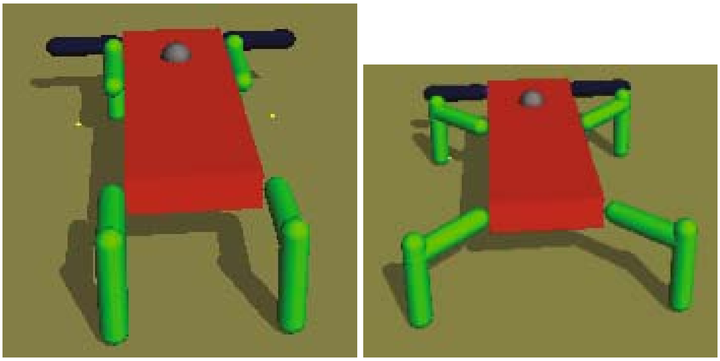

Fig. 5.1 The two virtual robots used in [3]: **Robot 1** (left), **Robot 2** (right).

of locomoting toward a distantly located object, grasping the object and then lifting the object onto their back. These robots each had 13 degrees of freedom and were actuated by a form of artificial neural network known as a continuous time recurrent neural network (CTRNN)[5]. For more details about the robots' morphologies and neural controllers please refer to [3].

In all cases training began in an environment where the target object was placed directly in front of the robot. Through a form of genetic algorithm (a hill climber) the CTRNN parameters were optimized until the robot was capable of grasping and lifting the target object. At this point the optimization process was paused and the environment was altered such that the target object was moved slightly further away from the robot. The optimization process then resumed until the robot was capable of reaching the target object at its new location, followed by grasping it and lifting it. After each such success, the process was paused, the environment was altered to make the task more challening, and then optimization was resumed. While this general process was the same for all experiments performed, what varied was the ways in which the target object was repositioned, known as the scaffolding schedule.

Specifically, four scaffolding schedules were investigated (see Fig. 5.2). The first scaffolding schedule, referred to as '**T**', placed the target object in front of the robot at increasing distances until the target object was a distance of three meters from the robot. It was observed that by this distance, the robot must have learned a stable gait to reach the target object. As distance was increased past three meters the target object was moved out in both directions along the line perpendicular to the robot's sagittal plane, requiring two sub-evaluations: one sub-evaluation with the target object placed in front and to the left, and another in which the target object is placed in front and to the right of the robot. This schedule forced the robot to learn forward locomotion with object manipulation followed by directed locomotion with object manipulation.

The second schedule used ('**C**') moved the target concurrently along the perimeter of circles with radius 5 meters and centers located at 5 and -5 meters with respect to the robot's initial position. In this case two sub-evaluations were always used. The final two schedules both moved the target object away from the robot linearly

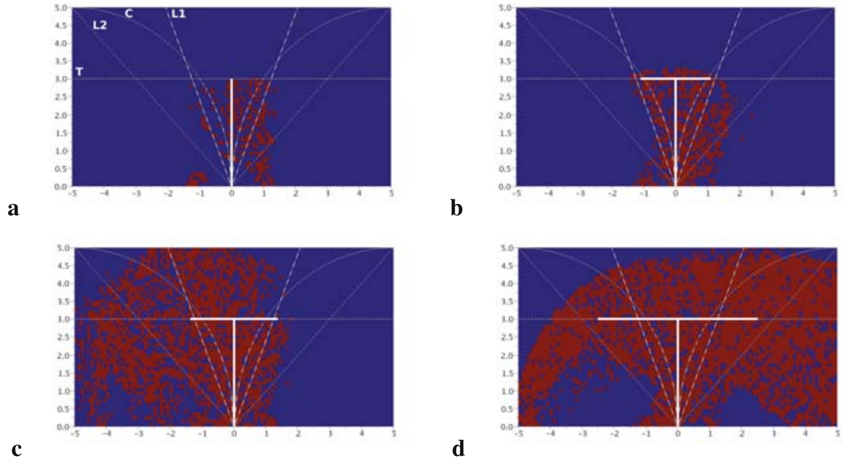

Fig. 5.2 Sample generalization plots from evolution of a generalized controller on robot 2 (red indicates the robot was successful at picking up the target object at that location) with the four scaffolding schedules superimposed. Specifically the plots shown are for controllers that were successful at distances of 3 meters (**a**), 3.2 meters (**b**), 3.3 meters (**c**) and 3.92 (**d**) the final training distance reached in this run.

on both sides. One did so with a slope $m = 1/\tan(22.5°)$ ('**L1**') and the other did so with a slope $m = 1/\tan(45°) = 1$ ('**L2**'). All three of these schedules, to varying degrees, forced the robot to learn to turn towards the target before or while learning locomotion.

After completion of a given training experiment two metrics were used to evaluate success. The first was the adaptation rate: how far from the robot the shaping algorithm moved the target object during training. Since the target object was only moved further away when a controller was found to be successful at the previous distance this metric gave an indication of how rapidly the robot could adapt to a changing environment. However, it did not measure how successful a given CTRNN would be in unseen environments. For this purpose a second metric was devised. Know as a generalization metric, this metric involved creating a grid extending from 5 meters left to 5 meters right of the robot's initial position and forward 5 meters, and systematically testing how well a given controller performed the task for a sampling of target object locations within this grid located at regular intervals. The fraction of these locations that the robot instantiated with this controller could succesfully complete the task would be the controller's generalization score.

Figure 5.3 depicts the mean and standard error score achieved on both of these metrics across 100 independent runs per robot per scaffolding schedule. Notice how the **T** scaffolding schedule significantly outperformed the other three schedules both in training distance achieved and generalization for both robots. Comparing performances between robots, it is noted that the **T** schedule evolved significantly more generalized controllers with the second robot (left hand grouping in Fig. 5.3b) while

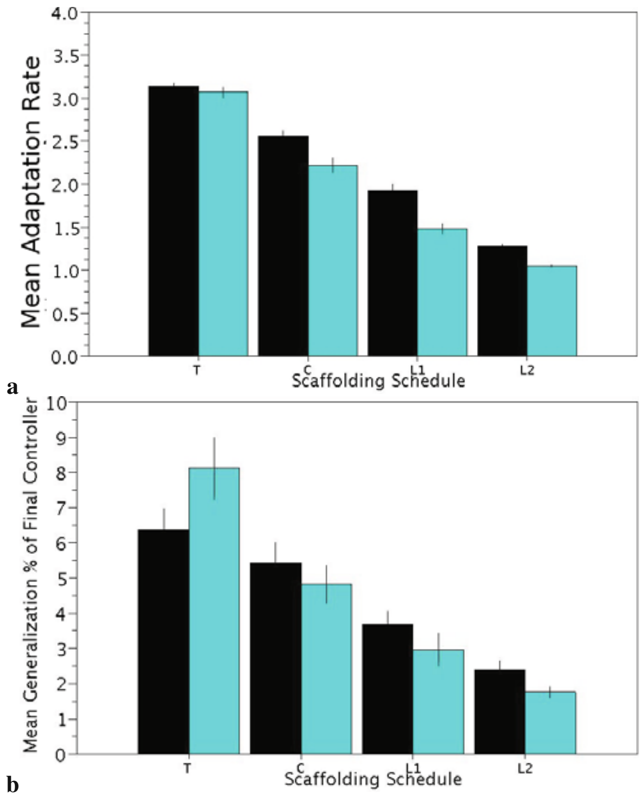

Fig. 5.3 Mean adaptation rate (**a**) and mean generalization % of final CTRNN (**b**) across the 100 runs for each of the two virtual robots (robot 1 in black, robot 2 in blue) and each of the four scaffolding schedules. All plots include standard error bars. Notice that while the mean generalization score for each set of runs was under 10% in all instances, there were runs in each set that found controllers with much higher generalization values. The generalization scores for the final controllers from the top five runs from each set are given in Table 5.1.

reaching similar final training distances as the first robot (left hand grouping in Fig. 5.3a). While the relative performance of the four schedules remained consistent across robots, the three other schedules led to slightly less generalized controllers with the second robot (three right hand groupings in Fig. 5.3b).

This means that both the morphology and training order are important for training a robot capable of completing the given task. Note that the schedules that pressured the robot to learn turning toward the target object either before or while learning to locomote were less successful than the one that pressured the robot to learn to locomote first. It therefore can be said that forward locomotion should be learned before turning, for both robot morphologies. As can be seen in Fig. 5.3 the probability of training a controller to enable taxis and object manipulation is inversely proportional to the pressure to learn turning before locomotion: the **T**, **C**, **L1**, and **L2** schedules

Table 5.1 Five best generalization values of final controllers from each set.

Schedule:	**T**	**C**	**L1**	**L2**
Robot 1:	53.6%	32.5%	23.3%	13.2%
	20.2%	28.3%	19.7%	12.7%
	16.6%	24.7%	14.9%	9.7%
	15.2%	24.3%	13.2%	9.2%
	15.1%	22.7%	11.5%	9.0%
Robot 2:	57.7%	26.3%	24.7%	12.6%
	40.4%	24.8%	24.1%	8.9%
	28.4%	21.4%	21.9%	7.6%
	27.4%	19.3%	19.6%	5.8%
	26.4%	19.1%	13.5%	4.8%

decline in performance, but increase in the pressure they exert to learn turning before locomotion.

This work demonstrated that with the proper scaffolding schedule (**T**) it is possible to evolve controllers capable of performing a non-trivial sequence of behaviors even in previously unseen environments. Moreover it demonstrated that altering morphology can impact the performance achievable through incremental shaping: robot 2 resulted in more generalized behaviors than robot 1. However, for the two morphologies considered the sequence in which behaviors should be learned remained the same. Robot 2's splayed legs made turning easier (see [3] for a discussion of this), however scaffolding schedules that selected for turning before locomotion was learned were not better able to integrate object manipulation, turning and locomotion into a controller using this body plan. Therefore it is concluded that the task environment, the learning algorithm, and/or the evolvability of CTRNNs dictate learning sequence more than morphology does.

More work remains to be done to strengthen this conclusion. Does this result hold across additional, uninvestigated, morphologies? How would evolving the robot's body plan along with its controller effect the sensitivity of the training procedure to the order in which behaviors are learned. The intuition is that evolving morphology would reduce this dependency and yield a more scalable method for realizing multiple dynamic behaviors in intelligent agents.

5.3 Specialization in a Morphologically Homogeneous Robot

Another recent publication [2] used a similar experimental framework as the work just discussed to investigate a different problem. In this case the research question was not about the order in which the behaviors should be learned, but about what variables influence the frequency of finding functionally specialized controllers – that is, controllers that devoted part of the robot's body (it's front legs) to a single behavior (object manipulation) rather than using that body part for multiple

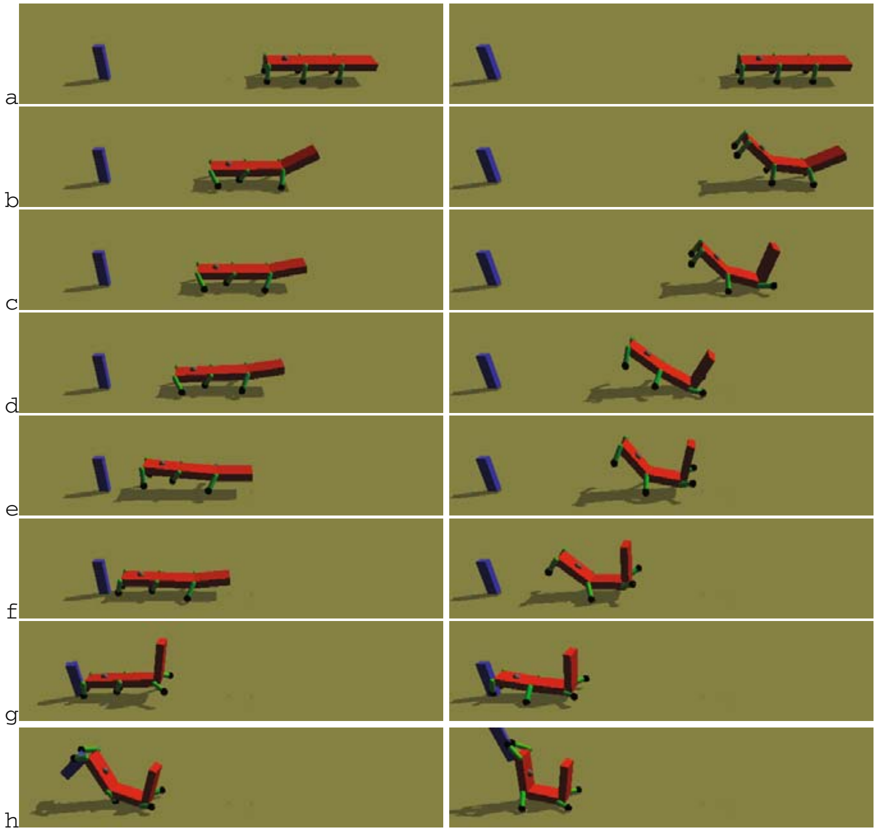

Fig. 5.4 Left Sample functionally generalized controller. This controller used the robot's front legs for propulsion during locomotion and for grasping and lifting of the target object. **Right** Sample functionally specialized controller. The robot's front body segment was raised and the front feet are kept off the ground during locomotion, i.e. they were only used for grasping the target object.

behaviors. Specifically the virtual robot investigated (Fig. 5.4) was a hexapod composed of three homogenous segments. It was designed such that the front segment could participate in locomotion and object manipulation (Fig. 5.4, left), or it may have become specialized such that it only participated in object manipulation (Fig. 5.4, right). In this way, selection pressure dictated the presence and degree of functional specialization rather than enforcing such specialization *a priori*.

This robot, like those described in the previous section, was trained with an incremental shaping algorithm coupled to a hill climber. Additionally, like those robots, this robot was controlled by a CTRNN. Several different experimental regimes were investigated with different initial environmental conditions and robot sensor configurations aimed at biasing the search process towards different solutions. For example, the first regime started with the robot's front segment rotated upwards 90° such that it was perpendicular to the ground with the front feet pointing forward and the target object initially placed directly in front of the robot. This configuration intuitively should have biased the evolutionary process towards finding controllers that specialized the front legs for grasping, since there was initially no evolutionary pressure for them to participate in locomotion, and indeed many of the runs from this regime found specialized controllers.

A second regime, conversely, started with the robot having all 6 feet on the ground. It was thought that this would bias the search toward controllers that did incorporate their front legs into their locomotion strategy, but this turned out to not be the case: a similar number of runs from this regime as compared to the first found controllers that specialized the front legs for object manipulation. An additional experiment began with the target object moved two meters in front of the robot, here it was thought that this would provide further bias towards incorporating the front legs into the locomotion strategy since, while learning to locomote initially there was no pressure for the front legs to be used for anything else, but once again a similar number of specialized controllers was found as compared with the previous two regimes.

Finally, a fourth regime with the same initial environmental conditions as the second regime, but with two additional sensors added to the robot and wired to its controller: joint angle sensors for the two joints connecting the body segments. The controllers that evolved in this regime not only performed better in the sense that they adapted more rapidly to changes in the target object's position during training as compared to the second regime, but also were more likely to be functionally specialized when compared to the other three regimes (blue bars in Fig. 5.5).

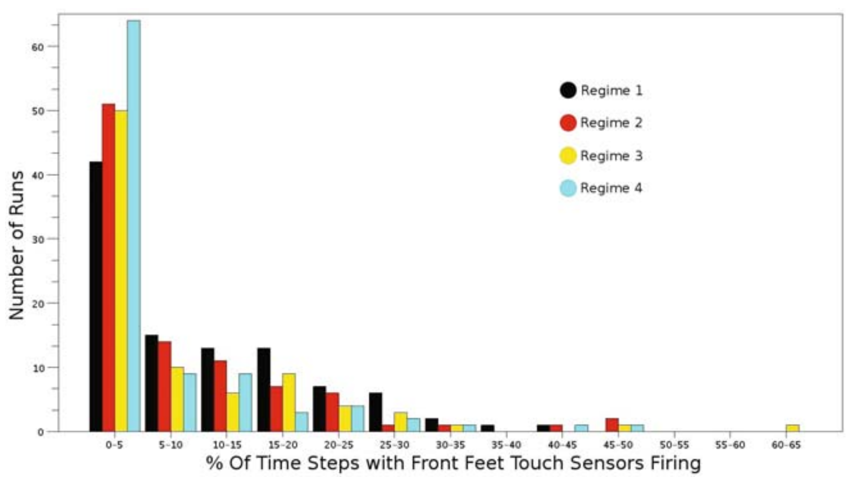

Fig. 5.5 Histogram of a specialization metric for each of the four regimes. All runs in which the target object reached at least three meters are included. See [2] for a description of this metric.

After noting that all four regimes were able to successfully learn both locomotion and object manipulation in the majority of trials the question arises as to why evolution tended to converge on functionally specialized behaviors, and why the inclusion of additional sensors caused an increase in the frequency of converging on such behaviors. Three possible hypotheses are: (1) functionally specialized controllers are more evolvable, and therefore supplanted less specialized controllers during an evolutionary run, (2) evolution initially discovered a specialized or generalized controller, and subsequently improved on that behavior but did not increase or decrease specialization, and (3) functionally specialized behaviors more easily allow for active perception [15].

While the first two hypothesis seem to be quite plausible, both were invalidated in [2]. The remainder of the space here will be spent discussing the more likely and potentially more interesting hypothesis number 3. According to that hypothesis, it may be that the robot was better able to actively perceive the proximity of the object—and therefore determine desirable conditions for lifting—if the front legs did not participate in locomotion, because then the touch sensors would only fire when in contact with the target object. Indeed, it has been demonstrated in the literature that active categorical perception may evolve in learning agents [4]. Moreover, providing the robot with additional proprioceptive feedback in regime 4 not only increased the prevalence of functional specialization (as shown in Fig. 5.5), but also the adaptation rate within those runs that produced specialized controllers. It is plausible that these added sensors allowed for better active perception as the touch sensors and sensed body posture may have together indicated appropriate conditions for object manipulation.

Several additional experiments were designed to test this hypothesis. These experiments followed the theme of the second and fourth regimes: fixing the initial environmental conditions but varying the sensors that the robot was provided with. It was demonstrated that adaptation rate declined as the included sensors provided less information in regards to desirable conditions for lifting. Specifically it was found that the main body joints were the most informative, while the front leg angles provided some information about the relative position of the front feet, but as the sensors are moved toward the rear of the body less of this relevant information would be available, and so the adaptation rate declined. This point was further demonstrated by an experiment that included joint angle sensors on every single joint. In this case the adaptation rate was not substantially improved compared with just including the most useful pair (those on the main body segments). Additionally it was shown in a further experiment that additional touch sensors improve performance even more so than any angle sensors do, because touch sensors provided the most direct evidence as to which feet are on the ground and/or touching the target object.

To verify that the additional sensors provided relevant information useful for the task and did not merely aid in locomotion, virtual robots were instantiated with the same sensor configurations and were evolved for locomotion alone. This consisted of expanding the range of the robot's distance sensors and placing the target object a large (100 m) distance away. Fitness was calculated as the fraction of distance

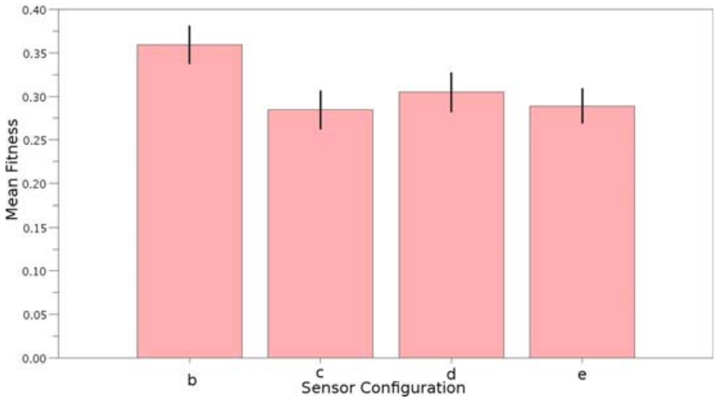

Fig. 5.6 Mean fitness with standard errors when selecting for just locomotion with four different pairs of joint angle sensors: **(b)** joint angle sensors on inter-segmental joints, **(c)** front leg joint angle sensors, **(d)** middle leg joint angle sensors, and **(e)** rear leg joint angle sensors.

between the start location and the target object location that the robot was able to cover in a set amount of time. Fig. 5.6 shows the mean fitnesses along with standard error bars from these experiments grouped by sensor configuration. Note that while including the joint angle sensors on the joints connecting the main body segments (**b**) led to improved locomotion performance, there was no significant difference between the performance of the other three sensor sets. This provided further evidence that the differences observed across these configurations above were due to active perception.

In conclusion, it was shown here that evolution can tune the amount of functional specialization of different parts of the body. It is predicted that if the morphology as well as the controller of the robot were under evolutionary control evolution would then specialize both the morphology and function for different body parts as the task environment dictates. Future work will test this prediction by evolving morphology as well as control. The hope is that this will prove to be a more fruitful method for realizing robots capable of an increasing number of behaviors, rather than fixing the body plan and manually assigning function to structure.

References

1. Anderson, M.: Embodied Cognition: A field guide. Artificial Intelligence 149(1), 91–130 (2003)
2. Auerbach, J., Bongard, J.C.: Evolution of functional specialization in a morphologically homogeneous robot. In: Proceedings of the Genetic and Evolutionary Computation Conference, GECCO (2009)
3. Auerbach, J., Bongard, J.C.: How robot morphology and training order affect the learning of multiple behaviors. In: Proceedings of the IEEE Congress on Evolutionary Computation, CEC (2009)

4. Beer, R.: The Dynamics of Active Categorical Perception in an Evolved Model Agent. Adaptive Behavior 11(4), 209 (2003)
5. Beer, R.D.: Parameter space structure of continuous-time recurrent neural networks. Neural Comp. 18(12), 3009–3051 (2006)
6. Beer, R.D.: The dynamics of brain-body-environment systems: A status report. In: Calvo, P., Gomila, A. (eds.) Handbook of Cognitive Science: An Embodied Approach, pp. 99–120. Elsevier, Amsterdam (2008)
7. Bongard, J.: Behavior chaining: incremental behavioral integration for evolutionary robotics. In: Bullock, S., Noble, J., Watson, R., Bedau, M.A. (eds.) Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems, pp. 64–71. MIT Press, Cambridge (2008)
8. Brooks, R.: A robust layered control system for a mobile robot. IEEE Journal of Robotics and Automation, [legacy, pre - 1988] 2(1), 14–23 (1986), `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1087032`
9. Brooks, R.: Cambrian intelligence. MIT Press, Cambridge (1999)
10. Calabretta, R., Nolfi, S., Parisi, D., Wagner, G.P.: Duplication of modules facilitates the evolution of functional specialization. Artificial Life 6(1), 69–84 (2000)
11. Dorigo, M., Colombetti, M.: Robot shaping: Developing situated agents through learning. Artificial Intelligence 70(2), 321–370 (1994)
12. Harvey, I., Husbands, P., Cliff, D., Thompson, A., Jakobi, N.: Evolutionary robotics: the sussex approach. Robotics and Autonomous Systems 20, 205–224 (1997)
13. Hornby, G., Takamura, S., Yamamoto, T., Fujita, M.: Autonomous evolution of dynamic gaits with two quadruped robots. IEEE Transactions on Robotics 21(3), 402–410 (2005), doi:10.1109/TRO.2004.839222
14. Izquierdo, E., Buhrmann, T.: Analysis of a dynamical recurrent neural network evolved for two qualitatively different tasks: walking and chemotaxis. In: Bullock, S., Noble, J., Watson, R., Bedau, M.A. (eds.) Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems, pp. 257–264. MIT Press, Cambridge (2008)
15. Noë, A.: Action in Perception. MIT Press, Cambridge (2004)
16. Nolfi, S., Floreano, D.: Evolutionary Robotics: The Biology,Intelligence,and Technology. MIT Press, Cambridge (2000)
17. Pfeifer, R., Bongard, J.: How the Body Shapes the Way We Think: A New View of Intelligence. MIT Press, Cambridge (2006)
18. Reil, T., Husbands, P.: Evolution of central pattern generators for bipedal walking in a real-time physics environment. IEEE Transactions on Evolutionary Computation 6(2), 159–168 (2002), doi:10.1109/4235.996015
19. Saksida, L., Raymond, S., Touretzky, D.S.: Shaping robot behavior using principles from instrumental conditioning. Robotics and Autonomous Systems 22, 231–249 (1997)
20. Singh, S.P.: Transfer of learning across sequential tasks. Machine Learning 8, 323–339 (1992)
21. Wood, D., Bruner, J., Ross, G.: The role of tutoring in problem solving. J. Child Psychol. Psychiatry 17(2), 89–100 (1976)