# Chaotic Exploration and Learning of Locomotion Behaviors

**Yoonsik Shim**
*necromax@gmail.com*
**Phil Husbands**
*p.husbands@sussex.ac.uk*
*Centre for Computational Neuroscience and Robotics, University of Sussex,*
*Falmer, Brighton BN1 9QG, U.K.*

We present a general and fully dynamic neural system, which exploits intrinsic chaotic dynamics, for the real-time goal-directed exploration and learning of the possible locomotion patterns of an articulated robot of an arbitrary morphology in an unknown environment. The controller is modeled as a network of neural oscillators that are initially coupled only through physical embodiment, and goal-directed exploration of coordinated motor patterns is achieved by chaotic search using adaptive bifurcation. The phase space of the indirectly coupled neural-body-environment system contains multiple transient or permanent self-organized dynamics, each of which is a candidate for a locomotion behavior. The adaptive bifurcation enables the system orbit to wander through various phase-coordinated states, using its intrinsic chaotic dynamics as a driving force, and stabilizes on to one of the states matching the given goal criteria. In order to improve the sustainability of useful transient patterns, sensory homeostasis has been introduced, which results in an increased diversity of motor outputs, thus achieving multiscale exploration. A rhythmic pattern discovered by this process is memorized and sustained by changing the wiring between initially disconnected oscillators using an adaptive synchronization method. Our results show that the novel neurorobotic system is able to create and learn multiple locomotion behaviors for a wide range of body configurations and physical environments and can readapt in realtime after sustaining damage.

## 1 Introduction

The possibility of exploiting intrinsic chaotic dynamics has recently attracted the attention of both neurobiologists interested in how animals learn to coordinate their limbs (Mpitsos, Burton, Creech, & Soinila, 1988; Kelso, 1995; Korn & Faure, 2003), for instance, in locomotion behaviors, and roboticists striving to develop better, more efficient locomotion systems for articulated autonomous robots (Kuniyoshi & Suzuki, 2004; Steingrube,

Timme, Worgötter, & Manoonpong, 2010). Chaotic dynamics emerging spontaneously from interactions of neural circuitry, bodies, and environments can be used to power a kind of search process as an embodied system explores its own possible motor behaviors. However, to date, it has not been clear how to harness chaos in a general goal-directed way such that desired adaptive sensorimotor behaviors can be explored, captured, and learned. We address this deficiency by presenting a general and fully dynamic embodied neural system, which exploits chaotic search through adaptive bifurcation, for the real-time goal-directed exploration and learning of the possible locomotion patterns of an articulated robot of an arbitrary morphology in an unknown environment.

Properly coordinated rhythmic movements for locomotion are ubiquitous in animals. Biological locomotor systems (usually involving coordinated limb movements) evolved to be highly adaptable, dextrous, and energy efficient. Consequently, they are a major source of inspiration when designing robot locomotion systems. Most biological locomotor systems involve neural networks acting as central pattern generators (CPGs), which are responsible for producing the basic rhythmic patterns for the oscillatory movement of limbs (Cohen, Rossignol, & Grillner, 1988; Stein, Grillner, Selverston, & Stuart, 1997). Understanding the subtleties of operation of such networks and how to design artificial versions for robotic applications are ongoing challenges (Ekeberg, 1993; Kimura, Akiyama, & Sakurama, 1999; Ijspeert, 2001; Ijspeert, Crespi, Ryczko, & Cabelguen, 2007).

While off-line search methods such as evolutionary algorithms or other global optimization processes have been extensively used to determine neural parameters for CPG-based robot locomotor systems (Gallagher, Beer, Espenschied, & Quinn, 1996; Ijspeert, 2001; Kamimura et al., 2003; Itoh, Taki, Kato, & Itoh, 2004; Floreano, Husbands, & Nolfi, 2008), the size and complexity of the search spaces often grow exponentially with regard to the number of variables, making the methods computationally expensive and time-consuming. Coupled with this, it is often very difficult to devise evaluation methods and metrics that can adequately cover the enormous number of unexpected situations that a robot can encounter during its lifetime, such as environmental change or body defects. This naturally led to efforts to develop adaptive methods that can be used online on the robot. Among these, reinforcement learning (RL) (Matsubara, Morimoto, Nakanishi, Sato, & Doya, 2006; Nakamura, Mori, Sato, & Ishii, 2007) and fast heuristic optimization algorithms (Sproewitz, Moeckel, Maye, & Ijspeert, 2008) have been successfully used. More systematic approaches such as continuous self-modeling, employing a number of stochastically optimized internal models (Bongard, Zykov, & Lipson, 2006), have also been developed. Although these are useful methods that allow more efficient online adaptation, they are not always free of the inherent difficulties of stochastic search (balancing exploration and exploitation, computational efficiency) and therefore often need to incorporate a priori knowledge or make use of

a biased learning strategy in order to simplify and speed up the learning process.

Partly because of these issues, the exploitation of intrinsic chaotic dynamics has recently emerged as an attractive alternative approach to the real-time online exploration of the space of embodied motor behaviors of a system. A number of bio-inspired robotics experiments have demonstrated its power in this context (Kuniyoshi & Suzuki, 2004; Pitti, Lungarella, & Kuniyoshi, 2005; Pitti, Niiyama, & Kuniyoshi, 2010). The research presented here significantly extends this direction by showing how to achieve an integrated system for the goal-directed exploration, capture, and learning of motor behaviors.

**1.1 Chaotic Neural Dynamics and Behavior.** A key influence on the current work is the growing body of observations of intrinsic chaotic dynamics in nervous systems (Guevara, Glass, Mackey, & Shrier, 1983; Rapp, Zimmerman, Albano, Deguzman, & Greenbaun, 1985; Freeman & Viana Di Prisco, 1986; Wright & Liley, 1996; Terman & Rubin, 2007). Some studies indicate intrinsic chaotic dynamics in animal motor behaviors at both the neural level (Rapp et al., 1985; Terman & Rubin, 2007) and the level of body and limb movement (Riley & Turvey, 2002). These seem particularly prevalent during developmental and learning phases (e.g., when learning to coordinate limbs) (Ohgi, Morita, Loo, & Mizuike, 2008). The existence of such dynamics in both normal and pathological brain states, at both global and microscopic scales (Wright & Liley, 1996), and in a variety of animals, supports the idea that chaos plays a fundamental role in neural mechanisms (Skarda & Freeman, 1987; Kuniyoshi & Sangawa, 2006).

Although the functional roles of chaotic dynamics in the nervous system are far from understood, a number of intriguing proposals have been put forward. Freeman and colleagues have hypothesized that chaotic background states in the rabbit olfactory system provide the system with "continued open-endedness and readiness to respond to completely novel as well as familiar input, without the requirements for an exhaustive memory search" (Skarda & Freeman, 1987). Kuniyoshi and Sangawa (2006) made the important suggestion that chaotic dynamics underpin crucial periods in animal development when brain-body-environment dynamics are explored in a spontaneous way as part of the process of acquiring motor skills.

Recent robotics studies have demonstrated that chaotic neural networks can indeed power the self-exploration of brain-body-environment dynamics in an embodied system, discovering stable patterns that can be incorporated into motor behaviors (Kuniyoshi & Suzuki, 2004; Kinjo, Nabeshima, Sangawa, & Kuniyoshi, 2008; Pitti et al., 2010).

**1.2 Embodiment and Locomotion.** Studying neural circuitry underlying the generation of rhythmic motor behavior in isolation ignores

the considerable advantage that can be obtained from incorporating the physical body and its environment—that is, exploiting the embodied nature of such behavior (Wheeler, 2005; Pfeifer & Bongard, 2007). In robotics, this has led to efforts to exploit ready-made functionality provided by the physical properties of an embodied system. One such line of inquiry involves using a frequency adaptive oscillator that can be tuned to the resonant frequency of the mechanical system (Buchli, Righetti, & Ijspeert, 2006; Raftery, Cusumano, & Sternad, 2008). Although this kind of adaptation accounts for some of the requirements for efficient locomotion, we believe that in general, the appropriate phase relationships between limbs should take priority when dealing with the creation of new motor patterns. One of the seminal works from this perspective is the exploration and acquisition of motor primitives, for a simple robot, using a mechanism that is embodied as a coupled chaotic field (Kuniyoshi & Suzuki, 2004; Pitti et al., 2005). That work modeled an extreme version of embodied coupling that had no electrical connection between neural units, with all neural coupling acting indirectly through body-environment dynamics. Neural oscillators were implemented using a simple logistic map with chaotic behavior, and the system dynamics rapidly developed to a stable, coherent rhythmic motion by using mutual entrainment between the neural circuit and the body-environment system. The process was completely deterministic, not making use of any random search method. More tractable systems (Pitti et al., 2010) have shown that a simple 2D simulated biped controlled by indirectly coupled chaotic maps can generate stable locomotion when the coupling strength between controller and body was set in the specific regime of phase synchronization. Phase synchronization between chaotic controller and physical system allows the flexible self-assembly of motor patterns and adaptive frequency matching to the resonant frequency of the body. However, the motor patterns that emerge through phase synchronization do not necessarily produce sustained locomotion behaviors unless the coupling strengths are properly set for a given neuromechanical system. Also, a more biologically plausible system was developed by Kuniyoshi and Sangawa (2006) in which a realistic musculo-skeletal model was employed with neural control circuits consisting of model CPGs. This was embedded within a larger system involving cortical maps. The biomechanical system was modeled as a series of redundant muscles acting on a joint, and information on the muscle combinations for any discovered coherent motor patterns was engraved on the model cortices as a sensorimotor representation. Later work (Kinjo et al., 2008) demonstrated the learning and replay of a motor pattern by adding a simple perceptron with a backpropagation learning on top of the previously learned sensorimotor maps. They showed that the representative power of the self-organized sensorimotor maps can greatly simplify the nontrivial sensorimotor learning problem into a simple mapping between the sensor and motor maps, but the learning pattern was manually fed to the system during learning; hence, it cannot be regarded

as an example of an autonomous and goal-directed exploration-learning scheme.

Until now, concrete general methodologies for applying such techniques to the automatic generation of desired motor patterns for autonomous robots have remained elusive. In this letter, we build on the essential concepts of prior work, extending and generalizing it as we attempt to develop a generally applicable methodology based around self-organization through chaotic dynamics for neural-body-environment coupled systems. We present a study of goal directed online exploration of rhythmic motor patterns in an oscillator system coupled through physical embodiment, specifically generating forward locomotion behaviors without prior knowledge of the body morphology or its physical environment. This is explored in the context of simulated limbed robots. In an important departure from the previous work outlined above, our recent study (Shim & Husbands, 2010) introduced an approach to explore and drive system dynamics toward a desired state by employing the concept of chaotic mode transition with external feedback (Davis, 1990), which exploits the intrinsic chaoticity of a system orbit as a perturbation force to explore multiple synchronized states of the system, and stabilizes the orbit by decreasing its chaoticity according to a feedback signal that evaluates the behavior. This enabled the system to perform a deterministic search guided by a global feedback signal from the physical system, which facilitates an active exploration toward a desired behavior. This preliminary work showed how to guide the system orbit to selectively settle in one of the stable patterns, but the system was restricted in that it was unable to capture and learn high-performing transient (unstable) patterns. The research described in this letter enhances our previous study by addressing those deficiencies and provides a coherent integration of these procedures into a dynamical systems framework, building a complete self-driven exploration-capture-learning system.

## 2  Chaoticity as a Perturbation Strength

Conventional optimization strategies generally use (external) stochastic perturbations on system parameters for search space exploration. However, a few studies have addressed the effectiveness of a chaotic system replacing a stochastic source (Parker & Chua, 1989; Ott, Sauer, & Yorke, 1994), and have found that a deterministic chaotic generator outperforms a stochastic random explorer (Zhang & Shao, 2001; Morihiro, Isokawa, Matsui, & Nishimura, 2005). In these cases, the chaotic dynamics acts as an external module generating perturbations that cause system parameters to wander in parameter space. However, the adaptive chaotic search method presented here, using bifurcation to chaos, can directly drive the phase orbit of a bodily coupled system (where the neural elements are coupled indirectly through physical embodiment) for exploration because of the endogenous existence of chaotic dynamics in the system itself. The intrinsic

dynamics of the system naturally power the search process without the need for external sources of noise.

The general idea of applying a chaotic search method that uses adaptive parametric feedback control had been previously presented in the field of optical sciences (Davis, 1990; Aida & Davis, 1994) and for memory search, where memory is stored as cyclical pattern sequences in a neural network (Nara & Davis, 1992). It has been argued that this method should be generally applicable when the target device is capable of supporting a variety of stable modes, between which there exist chaotic transitions, and which interacts with its environment such that there exists a feedback signal evaluating whether the mode is suitable or not. Chaotic transitions allow the system to try each of the modes sequentially, and the mode evaluated as suitable is selected and stabilized by changing a device chaoticity parameter to take it into a multistable regime. This can be thought of as a controlled version of the concept of chaotic itinerancy (Kaneko, 2003), where the system wanders from one quasi-attractor to another, getting entrained in each of them for a while. An indirectly coupled neural-body-environmental system, such as the one used in this letter, has the required characteristics of such a device, including multiple coordinated oscillation modes. It is known that a properly designed oscillator network can have multiple synchronized states that exhibit stable oscillations for both discrete (Feudel, Grebogi, & Yorke, 1996) and continuous (Vadivasova, Sosnovtseva, Balanov, & Astakhov, 1999) systems, and the structure of emergent behavior in these systems often reflects the spatial distribution of coupling strengths (Kaneko, 1994). Accordingly, a network of oscillators coupled through physical embodiment forms multiple synchronized states that reflect the body schema and its interactions with the environment, and each of them represents a potential candidate for meaningful locomotion behavior.

A conceptual description of the chaotic search process is illustrated in Figure 1. The goal of the system can be regarded as finding and becoming entrained in the basin of a particular attractor that has high performance (denoted by C) while escaping from the low-performing attractors (A and B) regardless of the initial point in the state-space. The idea is to open a new pathway that connects those isolated basins through the use of an additional dimension afforded by changing the system dynamics through tuning the chaoticity according to the evaluation signal. The orbit will visit and evaluate each of the attractors (A, B, C) systematically, yet chaotically, by adaptively varying the bifurcation parameter of the system according to a feedback signal until it reaches the basin of the desired attractor. The process can be interpreted as a continuous and deterministic version of trial-and-error search that exploits the intrinsic chaotic behavior of the system.

## 3  The Integrated Exploration-Learning System

The architecture of the neural part of the system developed in this letter is based on Kuniyoshi and Sangawa (2006) model, which is inspired by
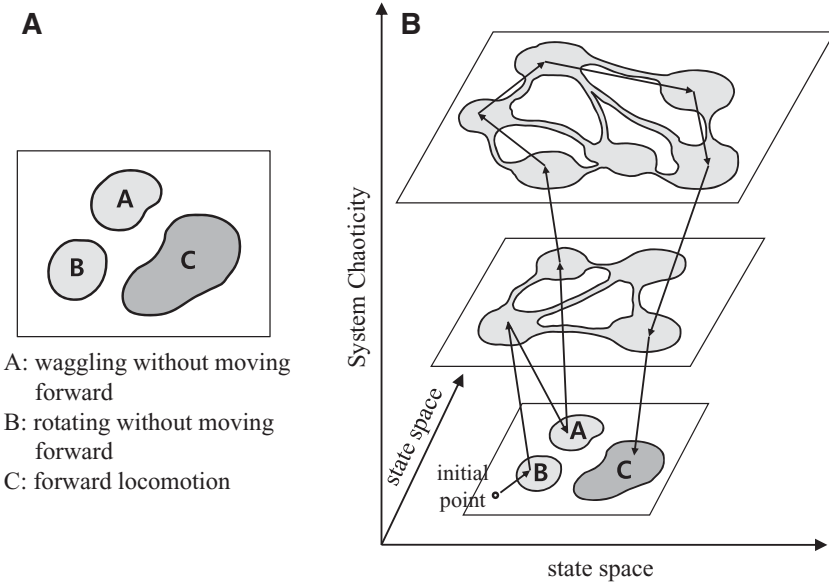
Figure 1: (A) A conceptual illustration of the state-space of a neural-body-environment system coupled through physical embodiment, which consists of three basins of attraction (A,B,C) with different performances. (B) An exploration process to find the desired attractor, C, by varying the complexity of the state-space landscape. Lump spaces and narrow passages in the landscapes of higher complexities represent quasi-attractors and itinerant pathways, respectively.

the organization of spinobulbar units in the vertebrate spinal cord and the medulla oblongata (the lower part of the brainstem, which mainly deals with autonomic, rhythmic, involuntary functions). But we use a more compact and modular configuration for each joint of the limbed robot and significantly extend the model to allow goal-directed exploration and learning. It is intended to be applicable to a wide range of robotic systems. The architecture consists of a number of identical control modules connected to each of the body parts. Each neuromuscular system for a joint that receives afferent sensory input and gives motor output to an antagonistic muscle pair can be encapsulated as a single motor unit, and the whole system consists of $N$ identical motor units where $N$ is the number of degrees of freedom of the robot (see Figure 2). Therefore, the system consists of uncoupled identical weakly forced limit cycle oscillators and a series of first-order leaky integrator equations. Prior work has demonstrated that uncoupled weakly forced oscillator systems can operate in stable modes (Kuniyoshi & Sangawa, 2006), and since our extensions to this work mainly involve elements based on stable first-order dynamics, it was possible to develop
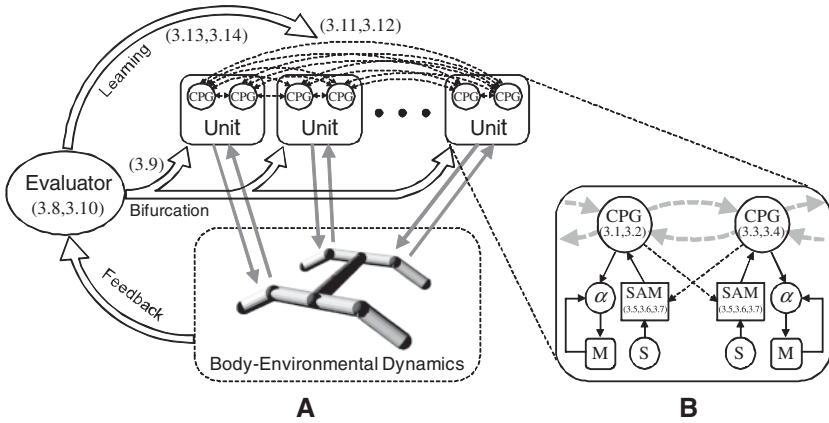
Figure 2: (A) An overview of the integrated exploration and learning scheme. Each degree of freedom in the robot's muscle joint system has a dedicated motor unit. Connections between the oscillator neurons are initially inactive, but they are weakly coupled through the body and environment. An evaluation feedback signal controls a global bifurcation parameter that alters the chaoticity of the CPGs. As the system stabilizes on a high-performing pattern, the bifurcation parameter reduces to zero and the connections between the oscillators become active, their weights being set by a learning procedure that is smoothly linked to the chaotic exploration process. The learning process further stabilizes, captures, and memorizes the motor patterns. (B) A motor unit for a single degree of freedom. A unit consists of two initially functionally disconnected oscillator (CPG) neurons, which receive integrated information of other oscillators in the system from the sensor (S), via body-environment interactions, and give a control signal to the muscle (M). SAM performs homeostatic adjustment of raw sensor signals by referring the antagonistic oscillator output, thus enhancing the synchronicity between the neural and physical system (see the text for further details). The numbers in parentheses refer to the corresponding equations used for each subcomponent.

a system that can be stably operated with an arbitrary body-environment configuration.

**3.1 The CPG Model.** Each motor unit has a pair of CPG neurons modeled by Bonhoeffer-van der Pol (BVP) equations (Asai, Nomura, Abe, & Sato, 2003), which drive the corresponding joint. When interacting with the body and environment, the motor unit can adjust its chaoticity by varying the difference between control parameters of the oscillators in the CPG pair. These differences change identically in all motor units as a function of the evaluation signal, acting as the global bifurcation parameter for the chaotic exploration with adaptive feedback. The BVP model allows the

phase relationship between CPG activity and body motion to be flexibly locked according to a loop delay (Ohgane, Ei, & Mahara, 2009), which is a beneficial feature for covering a range of sensorimotor delays originating from different body-environment configurations. All CPGs in the system are fully interconnected in the electrical sense, but they are functionally disconnected during exploration (by having zero connection weights). When the system dynamics are stabilized by discovering a useful pattern, the connection weights become nonzero, according to a learning procedure described later, and the fully interconnected network is activated.

In order to make them more amenable to this application, the center of rotation of each BVP oscillator was translated by $(0.2138, -0.7202)$ to locate it at the origin. The center of rotation was determined by averaging each pair of variables of a limit cycle for a sufficient duration. Including the sensory input and the coupling from other oscillators, a pair of oscillators in a motor unit $m$ is expressed as follows:

$$\tau \dot{x}_l = c \left( x_l - \frac{x_l^3}{3} - y_l + z_1 \right) + \delta (H_l(s_l) - x_l) + F_l^x, \tag{3.1}$$

$$\tau \dot{y}_l = \frac{1}{c} (x_l - b y_l + a) + \varepsilon H_l(s_l) + F_l^y, \tag{3.2}$$

$$\tau \dot{x}_r = c \left( x_r - \frac{x_r^3}{3} - y_r + z_2 \right) + \delta (H_r(s_r) - x_r) + F_r^x, \tag{3.3}$$

$$\tau \dot{y}_r = \frac{1}{c} (x_r - b y_r + a) + \varepsilon H_r(s_r) + F_r^y. \tag{3.4}$$

$\tau$ is a time constant, and $a = 0.7$, $b = 0.675$, $c = 1.75$ are the fixed parameters of the oscillator (Asai, Nomura, Abe et al., 2003). Each consecutive pair in the set of $2N$ oscillators is sequentially allocated to each motor unit as $l = 2m - 1$ and $r = 2m$ (we use expressions such as $^m x_l$ and $^m x_r$ to refer to the $m$th motor unit where it avoids confusion; see Figure 5). $\delta = 0.013$ and $\varepsilon = 0.022$ are the coupling strengths for afferent input $H(s)$, which is a function of raw sensor output $s$, processed by the sensor adaptation module (SAM) described in the next section. $F_i^j$ is a coupling term between oscillators and is subject to the learning process. $z_1$ and $z_2$ are the control parameters for adjusting the chaoticity of the motor unit. Their difference ($\mu = z_2 - z_1$) changes identically in all motor units and acts as the global bifurcation parameter. In the stable regime where the two control parameters are symmetric, it had been found (Asai, Nomura, Sato et al., 2003) that the two coupled BVP equations exhibit bistable phase locking of their oscillations in a parameter range of $0.6 < z_1 = z_2 < 0.88$. From the observation of a number of experiments on the oscillator dynamics, we chose to fix $z_2 = 0.73$ and to vary $z_1$ in order to ensure a higher probability of multistability of

the system in its stable regime. Note that we need to preserve the topology of indirect couplings between oscillators close to that of Asai's basic form (couplings from excitatory nodes to all nodes; Asai, Nomura, Abe et al., 2003), but slight variations in the sensor input term need be made for some sensor designs (refer to section A.2.1 in the appendix for examples).

**3.2 Homeostatic Sensory Regulation.** The sensor adaptation module (SAM) performs homeostatic adaptation (Turrigiano & Nelson, 2004; Turrigiano, 2008) for sensor input by calibrating the raw sensor signal using a linear transformation, which continuously adjusts the amplitude and offset of the periodic sensor signal in order to closely match its waveform to that of an antagonistic oscillator output. The sensory signal (in most cases, mechanosensory information from haptic sensors or muscle afferents) may vary according to the choice of sensors and the different body-environment interaction conditions. If the incoming signal is too large, the chaoticity of the system will be lost; if too small, the neural signals will be uncorrelated. The regulation of sensory activation ensures that the oscillator pair in a motor unit maintains a certain level of information exchange close to that of a weakly coupled oscillator pair so that the network dynamics are regulated within an appropriate range to generate flexible yet correlated activities. This also ensures the chaoticity of a motor unit is controlled in a systematic and collective way by the feedback signal regardless of the physical properties of the robotic system and the type of sensors.

The sensor processing function $H(s, t)$ is the implementation of a SAM. Given raw sensor signal $s$ and antagonistic oscillator output $n$, the adaptation function $H(s, t)$ is

$$H(s, t) = (s - \bar{s})e^{A(t)} + (\bar{s} + B(t)), \tag{3.5}$$

$$\tau_h \frac{dA(t)}{dt} = \sqrt{\overline{(n - \bar{n})^2}} - \sqrt{\overline{(H(s, t) - \overline{H(s, t)})^2}}, \tag{3.6}$$

$$\tau_h \frac{dB(t)}{dt} = \bar{n} - \overline{H(s, t)}, \tag{3.7}$$

where $\bar{x}$ represents the continuous running average of $x$ as calculated from $\tau_h d\bar{x}/dt = -\bar{x} + x$ (this meaning for $\bar{x}$ is used throughout the letter). The raw sensor signal $s$ is linearly transformed by a multiplicative factor $e^{A(t)}$ and an additive factor $B(t)$. The multiplicative function $A(t)$ is updated by comparing the difference of the root mean square of the temporal average of the squares of the antagonistic neural output $n$ and the transformed incoming signal $H(s, t)$, which is analogous to the signal energy that reflects the strength or amplitude. $B(t)$ is used as part of the scheme to remove the offset bias: each signal is subtracted by its average offset ($\bar{n}$ and $\overline{H(s, t)}$) before calculating the energy difference. $B(t)$ is updated by the offset difference between two signals. The timescale of adaptation should be set longer

than that of the oscillator, and we used $\tau_h$ as the timescale of performance evaluation ($\tau_E$) throughout this work, as described in the next section.

**3.3 Evaluation and Feedback Bifurcation.** During exploration, the bifurcation parameter continuously drives the system between stable and chaotic regimes as a function of the evaluation signal. The evaluation signal is determined by a ratio of the actual performance (e.g., forward speed) to the desired performance. If the performance reaches the desired performance, the bifurcation parameter decreases to zero, and the system stabilizes. Since the robotic system is arbitrary, we do not have prior knowledge of what level of performance it can achieve. Drawing on concepts from goal-setting strategies (Barlas & Yasarcan, 2006) and the Rescorla-Wagner model of conditioning (Rescorla & Wagner, 1972), the dynamics of the desired performance are modeled as a temporal average of the actual performance, such that the expectation of a desired goal is influenced by the history of the actual performance experienced.

In the experiments described next, the locomotion performance $E$ is measured by the forward speed of the robot. Since the system has no prior knowledge of the body morphology of the robot, it does not have direct access to the direction of movement or information on body orientation. In order to facilitate steady movement in one direction without gyrating in a small radius, the center of mass velocity of a robot was continuously averaged over a certain time window, and its magnitude was used as the performance of the system. The performance signal $E$ at any time instance can be calculated by applying a leaky integrator equation to the velocity vector as follows:

$$E(t) = \|\bar{\mathbf{v}}\|, \quad \tau_E \frac{d\bar{\mathbf{v}}}{dt} = -\bar{\mathbf{v}} + \mathbf{v}. \tag{3.8}$$

The timescale of integration was set as $\tau_E = 5T$ where $T$ ($\approx 8\tau$ in our BVP model) is the period of an oscillator. The time course of the bifurcation parameter $\mu$ ($= z_2 - z_1$) is given by

$$\tau_\mu \frac{d\mu}{dt} = -\mu + \mu_c G(E/E_d), \quad G(x) = 1/(1 + e^{16x-8}). \tag{3.9}$$

$\tau_\mu$ determines the timescale of the change of $\mu$ and is normally set faster ($\tau_\mu < T$) than the oscillation period ($T$) of the controller. If its value is too high, stabilization of the system dynamics is significantly delayed, which results in a partition mismatch (Aida & Davis, 1994). If it is too low, $\mu$ fluctuates too much according to the undulation of the robot movement, which acts as a disturbance for stabilization, or the system can become locked in a ring of undesirable patterns in a regime of intermediate chaoticity. $\tau_\mu = T$ was used throughout this work. $G(x)$ implements a decreasing sigmoid

function that maps monotonically from $(0, 1)$ to $(1, 0)$. $16x - 8$ shapes the sigmoid function so that the boundary value at $x = 1$ and its derivative ($[\frac{dG(x)}{dx}]_{x=1}$) become almost 0 so as to make the function smoothly vanish to zero. We automatically set $G(x) = 0$ when $x \geq 1$, since the bifurcation parameter $\mu$ should be zero in order to make the system completely stable. The dynamics of the desired locomotion performance, $E_d$, which slowly decays toward the current performance, is described by

$$\tau_d \frac{dE_d}{dt} = -E_d + E, \tag{3.10}$$

where $\tau_d$ is set sufficiently large so that $E_d$ does not follow $E$ too fast ($\tau_d = 25T$ in this work). Since $E_d$ continuously decays toward $E$, the changing speed of the control parameter $\mu$ depends on both $\tau_\mu$ and $\tau_d$. Since $G(x)$ decreases to zero asymptotically, $\mu$ was set to zero when it fell below a small threshold ($\approx 0.0001$), which also allows some margin for the system to stay in the stable regime ($\mu = 0$) despite the small oscillation of $E/E_d$ near unity. $\mu$ varies in the range $[0, \mu_c]$ where $\mu_c$ is the maximum level of chaoticity of the system. From the analysis of a single BVP oscillator, it is well known that it exhibits Hopf bifurcation with an increase of the parameter $z$ (Nomura, Sato, Doi, Segundo, & Stiber, 1993). An analytically estimated critical value of $z_1$ for equations 3.1 and 3.2, without their coupling and input terms, is $z_1 = z_c = 0.38247$, which indicates that the maximum possible value of $\mu_c$ is $\mu_{c(\max)} = 0.73 - z_c = 0.34753$. However, because the situation is different from the dynamics of a single oscillator, experiments on the robotic systems presented here revealed that the actual behavioral criticality of $\mu_c$ varies slightly (e.g., $\mu_{c(\max)} \approx 0.38$ for a swimmer, $\mu_{c(\max)} \approx 0.35$ for a quadruped) among different body and environmental settings. One way to determine the system-specific criticality of the control parameter is to simply observe the dynamics of the system with fixed $\mu = \mu_c$. If the system is beyond its critical state, one of the oscillators in the motor unit will generate near-zero amplitude by crossing a Hopf bifurcation point. Normally we chose $\mu_c$ to be slightly less than its maximum observed value, taking into consideration the saturating region of the sigmoidal function $G(x)$, so that it does not stay near the critical value for an unnecessarily long time when the oscillation amplitude becomes small.

   Although this evaluation strategy does not explicitly impose a bias for continuously striving for higher-performing behaviors (because of the dynamics of $E_d$), an implicit bias toward better-performing behaviors is partially imposed on the system by the way in which the bifurcation parameter $\mu$ behaves as a function of $E/E_d$ (see equation 3.9). Once the system has been stabilized to some behavior, the speed of system destabilization, for a given amount of behavior degeneration, depends on the performance level of the initially stabilized behavior. In the quasi-periodic regime that occupies a large portion of the entire system dynamics ($\mu$ in the lower saturation part

near zero and middle part of the sigmoid function, $G(x)$) the phase relationships of ongoing patterns shift slowly, while fast and catastrophic change occurs in the chaotic regime where $\mu$ is located around the upper saturation part (near $\mu_c$) of $G(x)$. When the actual performance $E$ of a stabilized behavior decreases by a given amount, a low-performing behavior is destroyed more quickly because $E_d$ will be relatively small, while a high-performing behavior is smoothly degenerated, giving it much more of a chance of being sustained or reentrained to itself. In this way, in practice, the system fully stabilizes onto behaviors that exhibit stable relatively high performance.

**3.4 Learning of Emergent Patterns.** As the exploration process stabilizes the system by discovering a high-performing locomotor behavior, the synaptic connections between oscillators are dynamically wired using an adaptive synchronization learning scheme. We adapted a learning model developed by Doya and Yoshizawa (1992), that decomposes the problem of weight learning between oscillators into a collection of cellular-wise processes by adjusting the input connection weights (also called the phase-lock matrix) of individual neurons to maintain a given phase relationship between the cellular activity and incoming signals. This is available only when the phase relationship between the neuronal activity and input signals is presented in advance, which provides a suitable interface for our exploration system. The coupling strengths are continually adjusted to follow the emergent patterns in parallel with the exploration process until the system is stabilized by discovering a desired pattern. When a switching parameter ($\eta$ in equation 3.14, which is determined by the global bifurcation parameter, $\mu$) is triggered around the onset of system stabilization, the decrease of the learning rate of the phase-lock matrix and the activation of oscillator couplings simultaneously take effect. The learning rules are set up such that during the exploration phase, the couplings effectively remain functionally inactive. As dictated by equations 3.13 and 3.14, the coupling gain $g$ is turned on when the bifurcation parameter $\mu$ goes to zero, which means learning is activated when the system is stabilized to some discovered pattern. Otherwise ($\mu \neq 0$) the system is in an exploration phase and $g$ is set to zero, which turns off the learning. Since the coupling is not strong and is activated gradually, highly unstable patterns that show short-lived high performance are naturally filtered by the instability of the pattern itself during the activation period (the system destabilizes and returns to the exploration phase). Thus, exploration and learning are merged as a continuous dynamical process such that the desired locomotion pattern is spontaneously explored, discovered, and memorized in a coherent way.

The oscillator learning process proceeds as follows. In order to improve the readability of the following equations, let us denote the states $x$ and $y$ (in equation 3.1 and equation 3.2) of oscillator $i$ as $x_i^1$ and $x_i^2$. Considering $M$ (=2N, where $N$ is the number of degrees of freedom of the robot) fully connected oscillators, the coupling term $F$ for state $j$ ($= 1, 2$) of oscillator $i$

$(x_i^j)$ can be written as

$$F_i^j = g \sum_{k=1}^{M} \sum_{l=1}^{2} p_{ik}^{jl} x_k^l \quad i \neq k, \tag{3.11}$$

$$\frac{dp_{ik}^{jl}}{dt} = \gamma \left\{ (x_i^j - \overline{x_i^j}) - \sum_{r=1}^{M} \sum_{s=1}^{2} p_{ir}^{js} (x_r^s - \overline{x_r^s}) \right\} (x_k^l - \overline{x_k^l}), \tag{3.12}$$

where $g$ is a small feedback gain term and $gp_{ik}^{jl}$ represents the adaptive connection strength coupling from $x_k^l$ to $x_i^j$, which forms a covariance-like learning rule. $\overline{x}$ is the continuous running average of $x$ calculated with time constant $\tau_E$. The full derivation of the learning rule can be found in section A.1. During the exploration process, the feedback gain $g$ and the weight learning rate $\gamma$ are adaptively adjusted according to the global control parameter $\mu$ so that the couplings between oscillators are gradually activated around the onset of system stabilization. $g$ and $\gamma$ are controlled according to

$$g = \alpha (1 - \eta), \quad \gamma = \beta \eta, \tag{3.13}$$

$$\tau_\eta \frac{d\eta}{dt} = -\eta + D(\mu - \epsilon), \tag{3.14}$$

where $\alpha$ and $\beta$ are constants and $D(x)$ is the heaviside function with very small $\epsilon = 0.0001$. As the incoming weights are learned in order to match the sum of afferent signals close to the oscillator's signal, it is sufficient to use $\alpha = \sqrt{\delta^2 + \varepsilon^2}$ (input weights in equations 3.1 and 3.2), which has similar intensity to the sensory input. $\beta = 1/\tau_E$ was set to have the same timescale as the evaluator. $\eta$ is the smooth activation signal that controls both the learning rate of connection weights and feedback gain according to the value of $\mu$. This signal gradually activates the functionally connected network rather than suddenly switching it on, thus preventing the destruction of stable patterns while allowing unstable ones to be filtered out.

## 4 Experiments with Simulated Robots

Detailed experiments with the framework described above used the two simulated robots shown in Figure 3: a four-armed aquatic swimmer and a quadruped.[1] Initial experiments used the swimmer, which has four fins,

---

[1]The flash streaming (FLV) as well as downloadable AVI files of the movies in this work are available online at http://www.informatics.sussex.ac.uk/research/groups/ccnr/movies/yssmovie.html. Videos 1–10 show the behaviors of 4-fin Swimmer and quadruped, and videos 11–15 show other kind of robots that use controllers identical to the quadruped's except the number of motor units (video 14).
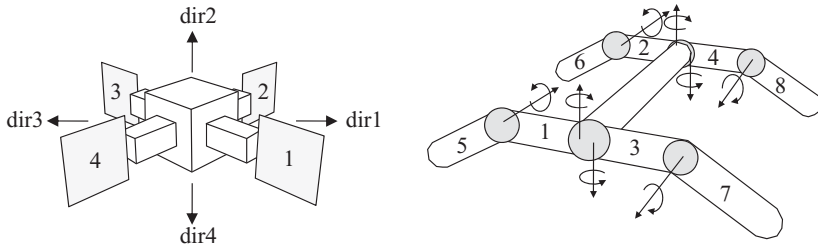
Figure 3: Robotic simulation models of a 4-fin swimmer (4-DOFs, 2D movement) and a quadruped (8-DOFs, 3D movement).

each at the end of a separate arm, placed in a simulated hydrodynamic planar (2D) environment. Since the information transfer between CPGs is mediated by sensory information, the information structure provided by physical embodiment is considerably influenced by the design and choice of sensory systems. While it is possible to use composite sensory information from multiple sensors (e.g., a combination of the input from fin sensors and muscle receptors), for simplicity we use only a single fin angle sensor for a motor unit. This requires a slightly modified sensor input term in the CPG equations in order to make the pair of CPGs in a motor unit deal with a single sensor (see equations A.11 and A.12). The functional structure of coupling between motor units through embodiment is formed by the transmission of hydraulic reaction forces from one arm to the others as the body articulates. The robot's radially symmetric shape in a 2D underwater environment is interesting because it makes generating continuous asymmetric propulsion forces challenging: forward locomotion is nontrivial. The robot will not be able to move in a single direction unless the movements of all four arms are successfully coordinated with appropriate phase differences.

**4.1 Exploration of Stable Patterns Without Oscillator Learning.** First, we fixed the bifurcation control parameter to the stable regime ($\mu = 0$, no chaotic search) and ran the 4-fin swimmer simulation to see what kinds of behaviors emerged from various initial states. More than 1000 simulation runs were tested in order to observe and categorize the behaviors. Basic movement behaviors of the swimmer were categorized into motion in four directions (along the body axes dir1, dir2, dir3, and dir4, as shown in Figure 3), which met expectations given the symmetric shape of the swimmer.

Taking the directional symmetry into account, we observed six different behaviors and classified them according to the locomotion performance, as shown in Table 1; their phase relationships are shown in Figure 4. The forward locomotion involves straight movements (ST), moving in circles (STC), and peg-leg (PL) motions. ST locomotion is a frog-like swimming action that has the highest performance (see Figure 8), and STC motion moves

Table 1: Categorized Emergent Behaviors Existing in the Stable Regime for the 4-Fin Swimmer.

| Pattern | Number of Variations | Average $E$ |
|---|---|---|
| 1. Straight (ST) | 4 (each dir) | 0.7 |
| 2. Circular (STC) | 8 (4×(CW,CCW)) | 0.6 |
| 3. Rotate (R) | 2 (CW,CCW) | 0.06 |
| 4. Peg-leg (PL) | 4 (each arm) | 0.04 |
| 5. Vibration (VB) | 2 (dir 1-3 and 2-4) | 0.03 |
| 6. Bound antiphase (BA) | 1 | 0.0 |
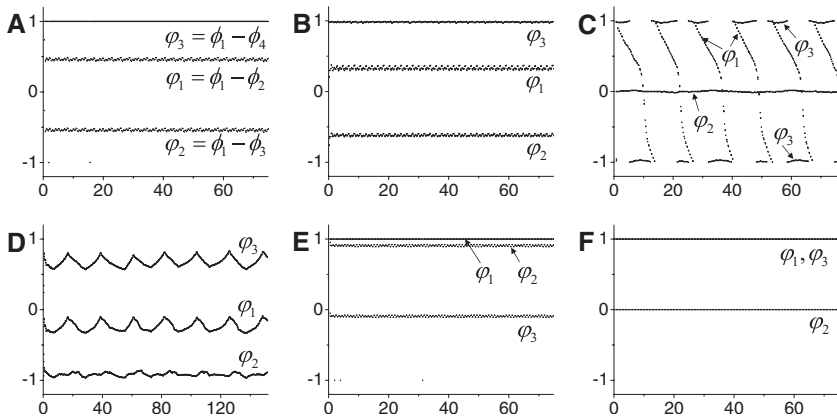


Figure 4: Behaviors of the 4-fin swimmer. Time (oscillator cycle) versus instantaneous phase differences ($\pi$ rad) for each behavior is depicted (calculated using Hilbert-Huang method; Huang et al., 1998): (A) ST dir2, (B) STC dir2 CCW, (C) PL arm2, (D) Rotate CW (period was doubled for clarity), (E) VB dir 1-3, and (F) BA. STC motions are similar to ST with slight variations of both phase and arm amplitude. The phase relationship of VB is qualitatively the same as BA with a different arm combination. The interlimb coordination can fluctuate (D) or even show continuous shift (C) when the corresponding arm moves irregularly with small amplitude; however, their qualitative behaviors are maintained by the body-environmental coupling.

in a circle due to a slight asymmetry between contralateral arms caused by passive fin dynamics and can be either clockwise or counterclockwise. PL motions involve one of the arms moving with a small amplitude while the other three arms all use the same large amplitude. The phase relationship of the PL pattern is essentially similar to that of bound antiphase, except that the amplitude of one arm is smaller than the others and its phase

continuously shifts (with a small irregularity) compared to the others, which achieves a slow forward locomotion by asymmetric propulsion forces.

Also nonlocomotion movements were observed such as bound antiphase (BA), vibration (VB), and rotation (R). BA motion results in no net movement of the robot torso due to antiphase locking between adjacent pairs of arms. VB arm movements are contralaterally antiphase and ipsilaterally in-phase based on the vibrating axis. The movements of arms in the rotation motion are out of phase with each other and fluctuate irregularly. The fluctuation and shifting of phase relationships suggest that an emergent behavior does not necessarily exhibit concrete phase locking between subsystems in the neuro-body-environment setting.

If a robot behavior was observed as being permanently sustaining, it was identified as an individual behavior. The number of completely stable behaviors in the absence of oscillator learning was determined to be six, without counting their variations. The shape of the 4-fin swimmer robot is radially symmetric, so different synchronized pairs of joints (variations) can exist for a single behavior. For example, the straight swimming behavior has four different combinations of synchronized joint pairs, all of which show the same frog-like swimming behavior. As shown in Table 1, there are 21 different arm coordinations when including all variations. Careful viewing reveals that the circling movement (STC) can show slightly different circling radii resulting from small differences in passive fin tilting, but these are too small to be considered separate distinguishable behaviors. In order to keep the analysis clearer, these kinds of variations are not counted as different behaviors.

Note that the PL patterns appear as a stable pattern only when sensory homeostasis is present. Sensor adaptation makes the lame arm synchronize with the corresponding motor unit with a small amplitude, resulting in the partial loss of the phase correlation with the other arms as it transfers the inertial or hydrodynamic forces less strongly to them through physical embodiment. Again, the motion of the other three arms is coordinated in such a way that the net forces are transferred at a reduced rate to the lame arm. Therefore, the homeostatic regulation of sensory signal results in an opposing effect, which leads to the diversification of limb motion, that is, the multiple combinations of the amplitudes and offsets of limb motions can be explored and stabilized by sending the standardized sensory input signals to the neural controller (see Figure 9). In turn, different limb-wise oscillations may cause different interlimb coordination as well.

The stable dynamics of the system begin to fluctuate as $\mu$ increases, exhibiting a series of transient dynamics from quasiperiodicity to chaos (see Figure 5). In the higher chaotic regime, complex transitory dynamics similar to chaotic itinerancy occurs, which drives the system to briskly explore the phase space. To see the effect of chaotic search, the distribution of visits to each of the behaviors identified in Table 1 was investigated under the presence and absence of chaotic search. One hundred simulations were
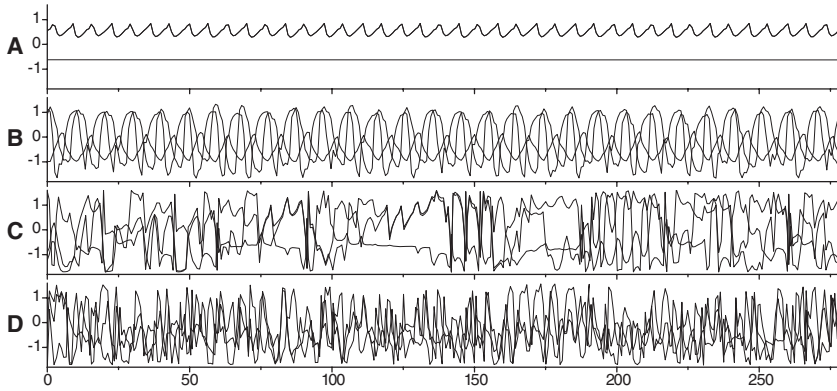
Figure 5: Time versus sampled trajectory of the neural outputs of the 4-fin swimmer with different $\mu$ fixed at (A) 0.2, (B) 0.32, (C) 0.336, and (D) 0.35. The agonist signals of three motor units ($^2x_l$, $^3x_l$, $^4x_l$) are plotted when $^1x_l$ crosses the singular axis (vertical axis located at the unstable equilibrium point) of the periodic orbit. The destabilized system with higher $\mu$ exhibits fully asynchronous (D) or intinerant (C) behaviors. Even synchronous dynamics (A, B) often possess small irregularities according to the initial condition and the physical environment.

performed for each case, and the visiting counts of six major behaviors were recorded by observation. Figure 6 shows a clear difference between the visiting ratios in the two cases, suggesting the effectiveness of chaotic search, which tended to settle on high-performing dynamically stable locomotion. During the search process, all variables and control parameters vary continuously as parts of the neuro-body-environment system, and the time evolution plots of phase differences, performances, and bifurcation parameter (see Figure 7A and 7B) show that the stabilization and destabilization of the system occur repeatedly in a trial-and-error manner until it settles on an effective form of locomotion. The sensor parameters (see Figures 7C and 7D) also change continuously and settle to different values through adaptation.

Due to the symmetric shape of the 4-fin swimmer, the BA motion has inherent dynamic stability with large basins in the phase space, so the system was often entrained in the BA pattern and sometimes took a relatively long time to reach one of the desired states. This deficiency, the so-called *deep-path* (Shim & Husbands, 2010), occurs when an orbit that tries to escape from BA by system destabilization is reinjected to BA, so the actual performance $E$ stays low and the desired performance $E_d$ decays close to $E$. This makes the time spent in the chaotic regime shorter, resulting in reduced exploration and increased time to escape. The escape orbit is often stabilized to PL patterns, which indicates that these patterns are located in the vicinity of BA in the phase space. However, the use of an adaptive $E_d$ (see equation 3.10),
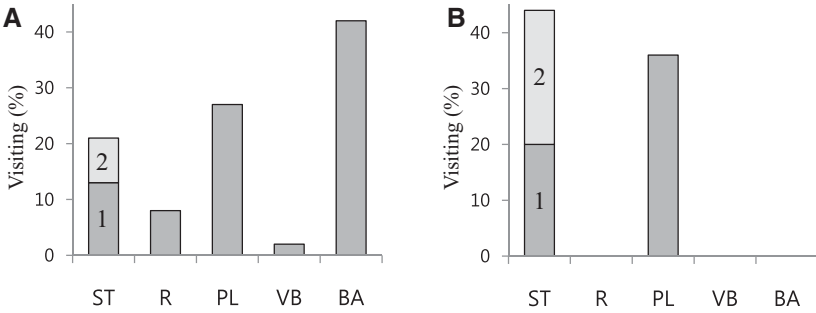
Figure 6: Visiting ratio of each pattern. Appearance of behaviors with (A) no control ($\mu = 0$) and (B) chaotic search. Every run was started from the random initial values of neural variables ([$-1, 1$]). The ST (number 1) and STC (number 2) movements are depicted as a stacked bar. Although PL movements have relatively low performances, they were frequently searched because of their high stability. The average performance of rotation (R) movement is slightly higher than that of PL, but because of the large fluctuations in body position without overall forward movement, it could not be stabilized by the search process.
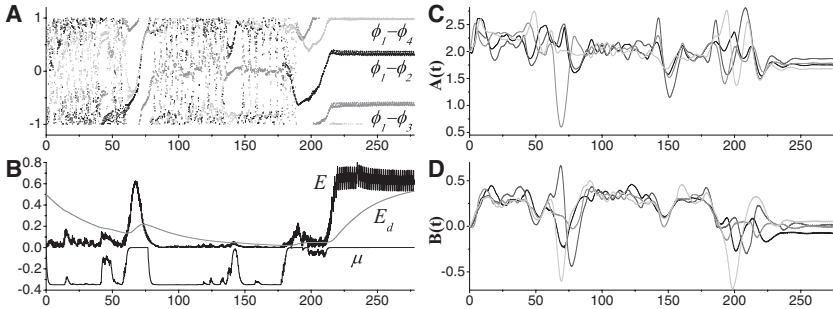


Figure 7: A typical example of exploration and stabilization from the swimmer experiment without learning (settles on to ST-dir2 locomotion). (A) Time versus instantaneous phase differences between arms 1-2, 1-3, and 1-4. (B) Performance, desired performance, and control parameter ($E$, $E_d$, and $\mu$ in equations 3.8 to 3.10). (C, D) Sensory adaptation parameters $A(t)$ and $B(t)$ (equation 3.5) for fins 1: black, 2: dark gray, 3: gray, and 4: light gray.

sensory adaptation and oscillator learning have all helped to significantly alleviate this issue. Figure 10 shows an example of the exploration time taken for stabilization of the systems with and without adaptation. The fixed sensor gain of the nonadaptive system was chosen to produce a simi-lar behavior category to the adaptive case. While the adaptive system was
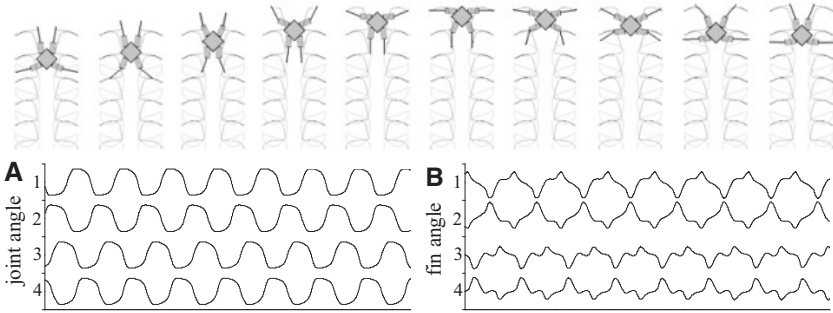
Figure 8: (Top) Snapshots of the straight swimming (ST dir3) behavior of the 4-fin swimmer. Images were taken every 1/10 gait cycle. The tip trajectories of the fore (fin 3, 4: black) and rear (fin 1, 2: gray) fins are shown. (Bottom) (A) Joint angles and (B) fin bending angles of the behavior. Each segment along the vertical axis indicates the range $[-1, 1]$ rad.
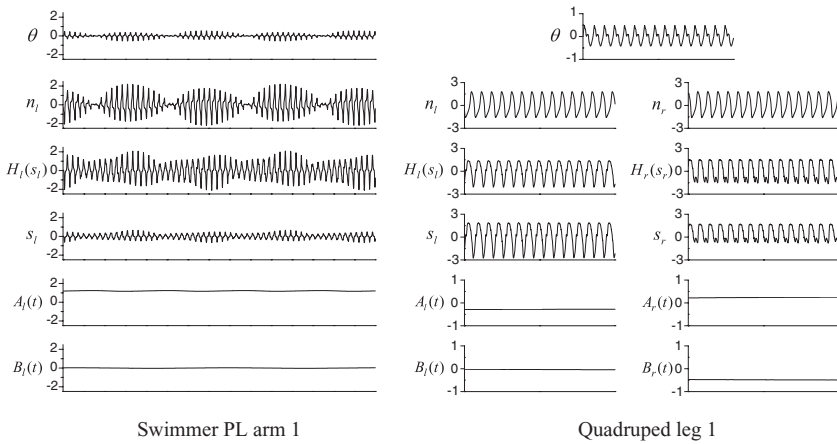


Figure 9: An example of the effect of sensory homeostasis. The behaviors of PL arm1 of the swimmer and the quadruped walking are depicted. From the top, each graph shows arm and leg angles ($\theta$), reference neural output ($n$), calibrated sensor input ($H(s)$), raw sensor signal ($s$), and the adaptation variables ($A(t)$) and $B(t)$). Since the swimmer uses a single sensor, only the behaviors of the agonist side (related to left CPGs) are shown, as the motions of the antagonist side are the same with different signs except for $A(s)$.

stabilized within 1000 cycles in general, a number of runs of the nonadaptive system showed it could take up to 10 times as long to stabilize compared to the adaptive system. Also the nonadaptive system exhibited bad-lock (Shim & Husbands, 2010) onto nonlocomotion patterns (rotation and vibration)
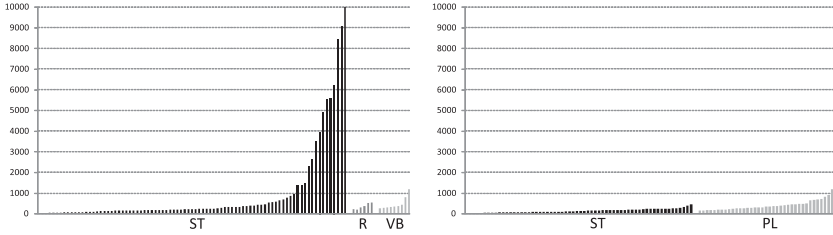
Figure 10: Final behaviors and their exploration time (oscillation cycles) until stabilization with (left) and without (right) sensory adaptation. For each case, 100 simulations were run, and each behavior was ordered by the time taken to stabilize. The raw sensor value for the system without adaptation used a fixed sensor function (see equation 3.5) of $H(s) = 7.4s$.
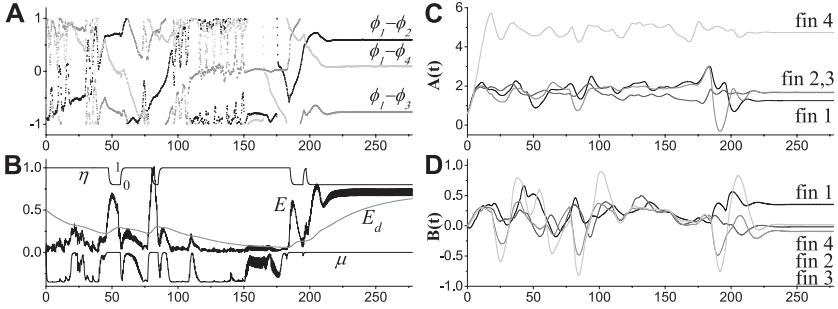


Figure 11: Exploration and capture of transient locomotor behavior in damaged-fin swimmer by oscillator learning. The oscillator wiring is triggered ($\eta$ in equation 3.13, 1: off, 0: on) at each stabilizing trial and attempts to sustain the performance of the discovered pattern. The sensor parameters $A(t)$ and $B(t)$ cooperate with the exploration learning for a given physical embodiment (see the text).

where the bifurcation parameter $\mu$ does not reach zero but oscillates near zero being phase-locked with other system variables.

**4.2 Stabilizing Transient Patterns by Oscillator Learning.** Often there are high-performing locomotion patterns that are not completely stable and appear for only a while during the exploration process. These transient target behaviors can be captured and memorized by the oscillator learning process. We tested this using a "damaged" version of the swimmer robot by reducing the length of one of its fins (damaged fin) or removing one of its arms (three-armed), such that there are few or no stable patterns in the stable regime but there exist a series of useful transient patterns.

Figure 11 shows the exploration and learning of the robot with a damaged fin, where the length of the fin on arm 4 was reduced by 90%. It had
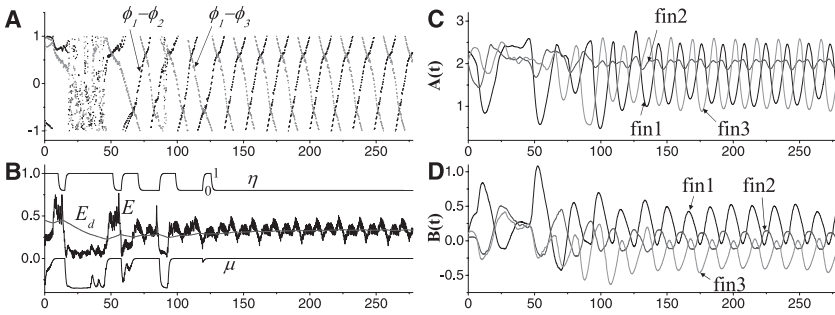
Figure 12: Alternating behavior of the 3-arm swimmer.

only one stable pattern whose phase relationship is the same as that of the BA pattern in the undamaged robot, which has almost zero performance. With learning, it captured one of the high-performing transient patterns after a few trials. The approximate direction of locomotion is toward dir-3. Figures 11C and 11D show that the sensor gain ($A(t)$) of the damaged fin (fin 4) was increased to amplify its signal, and the fact that fin 1 has the smallest gain tells us that arm 1 is the main source of propulsion. The salient deviation of the offset ($B(t)$) of the fin 1 sensor (opposite side of fins 2 and 3) indicates that the discovered transient pattern involves the oscillation of fin 1 in a tilted position, granted by its mechanical compliance; consequently it compensated the asymmetric hydrodynamic forces and achieved forward locomotion. The homeostatic sensory regulation participates in the exploration process as the slow variables diversify the course of transient patterns during search and slows them down at the onset of discovery, which is beneficial to the real-time pattern capture by oscillator learning. While the case of the swimmer robot has shown a relatively limited variety of patterns due to its strong, embodied coupling resulting from the densely structured physical environment it inhabits (the robot is always surrounded by liquid and hence is continually subjected to significant hydrodynamic forces), we will see later that the effect of sensory regulation on terrestrial movements becomes more prominent. Figure 12 shows a particular case of an alternative three-armed robot (formed by removing arm 4) where two different locomotion patterns are periodically exchanged while not losing the overall stability of the whole behavior. The robot alternates its moving direction between dir-3 and dir-4 by exchanging two unstable undulating motions. The periodicity of this conjoined behavior also exhibits a small degree of irregular fluctuation as in the case of loosely coordinated behaviors previously shown in Figure 4. However, being captured by oscillator coupling, it is sustained by global coordination between subsystems that include adaptive sensor dynamics.
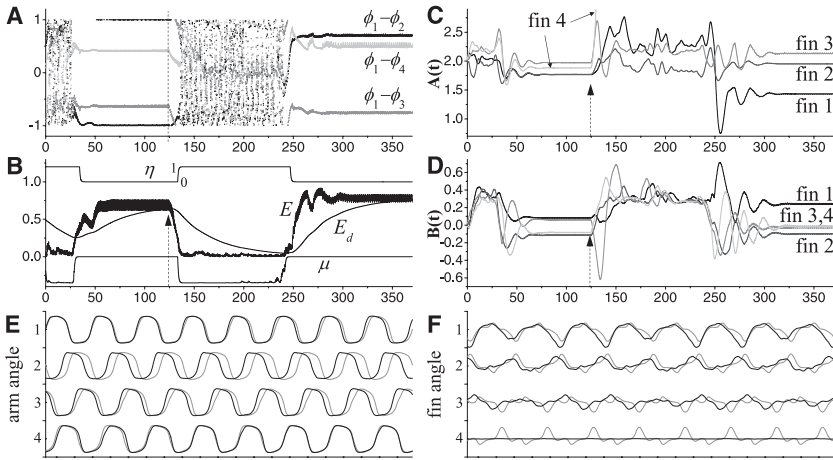
Figure 13: Real-time recovery after a radical change to the body (damage). (A) Instantaneous phase differences. (B) Performance variables and control parameters. Dashed lines and arrows indicate the time of damage, when the length of fin 4 is decreased to 1/10 of its original length. The sensor gain of (damaged) fin 4 ($A(t) \approx 5.0$) in panel C was truncated for a better view of the other gain plots. (E, F) The joint angles and the fin angles, respectively, where the undamaged motion (gray) and the readapted motion (black) are superposed. The fiducial point for the superposed plots was set to the starting point of arm angle 1 in panel E.

Since the oscillator learning process is automatically regulated by a control parameter ($\eta$), it is possible to operate the exploration-learning system continually without reset. Figure 13 shows a typical successful example of the real-time recovery of locomotion behavior after body damage of an unknown variety, that is, with no a priori knowledge. During an initially learned stable behavior (similar to STC-dir3-CCW), the same damage as in Figure 11 was sustained. The performance of the robot immediately dropped below $E_d$, and the system entered into the search phase. After a few hundred cycles, the system found a new locomotion behavior for the changed body (undulating movement similar to Figure 11). The superposed graphs of two behaviors (see Figures 13E and 13F) show a slight frequency increase in arm movements after recovery due to the change of mechanical impedance of the robot.

**4.3 Quadruped Locomotion.** We demonstrate the generality of the approach by also applying it to a quadruped robot in a 3D terrestrial environment. The stretching force (see equation A.13) experienced by a torsional muscle was used as the sensory signal and fed to the CPG in the relevant
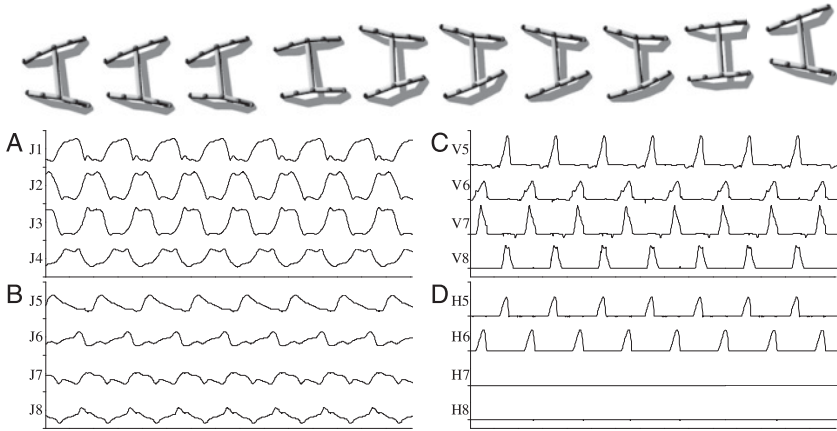
Figure 14: An example of a quadruped gait captured by the exploration-learning process. Snapshots were taken every 1/10 gait cycle. (A, B) The joint angles of limbs. (C) The horizontal speeds of each foot (the tips of limbs 5–8) in the direction of locomotion. (D) The height of each foot from the ground. The two rear feet (V7, V8, H7, H8) show stick-and-slip movements on the ground under Coulomb friction. The range of each plot is as follows; J1–J8: [−1.0, 1.0]rad; V5–V8: [0.0, 2.0]m/s; H5–H8: [0.0, 0.08]m.

motor unit. Under conditions where static stability against gravitational force is guaranteed in both the 2D swimmer and 3D quadruped, the walking machine has fewer behavioral constraints for producing forward locomotion since the resistance force is not always present in the 3D terrestrial environment (e.g., there is no friction on a leg as it moves through the air during a swing phase). The neural-body-environmental phase space of the quadruped can be envisaged as an undulating landscape of rolling hills, while the 2D swimmer case has a few deep basins of attraction. While this increased the number of candidate patterns for forward locomotion in the quadruped, there existed latent instabilities such as slipping due to dynamic friction or the spontaneous occurrence of sharp-amplitude, high-frequency perturbation stemming from the ground contact, all of which caused a slow degeneration of the ongoing locomotor pattern. In practice, the movement patterns of the quadruped observed in the stable regime of the oscillator system exhibited no ultimately permanently sustained behavior (also true in tests on other walking robots). Interestingly, locomotor patterns similar to the quadruped walking gait frequently emerged during exploration (see Figure 14). Other kinds of as-it-could-be gait patterns and their variations that exploit given active and passive dynamics were also observed, which are difficult to categorize qualitatively.

The degeneracy of locomotor behavior could be greatly improved by using homeostatic sensory adaptation and then completely stabilized by oscillator learning. Figures 15A and 15B show the system behaviors for quadrupeds with and without sensor adaptation. All experiments were started from the same initial condition. In Figure 15A, the sensor adaptation was turned off when the system was stabilized to the first discovered pattern. The performance of the emergent pattern in the adaptive system (see Figure 15B) degenerates much more slowly than in the nonadaptive case. Sensor adaptation prevented abrupt changes in phase relationships by buffering sudden changes of incoming sensor signals, so the initial movement pattern slowly changed, giving it a greater probability of being maintained. The patterns could be completely stabilized by introducing oscillator learning (see Figure 15C). However, if oscillator learning was presented without sensory adaptation (see Figure 15D), the pattern could not be sustained completely because the oscillator coupling was not strong enough to maintain the coordinated pattern against the degeneracy. As a result, the role of homeostatic sensory adaptation becomes more prominent in the case of terrestrial behaviors. The experiments with the 2D swimmer have shown little variance of sensor parameters after convergence, and pattern degeneracy was hardly observed, which indicates that the transient patterns of the swimmer are strongly attracted to a small number of stable patterns. The adaptation of the sensor parameters of the quadruped yielded more diverse values, where the offset parameter ($B(t)$) of lower leg muscles (leg 5–8) typically showed notable deviation under the effect of constant body weight. In a few cases, the speed of degeneracy under the control of the oscillators after adaptation is so slow that the locomotor behavior, which appears stable, is eventually destroyed after a very long period of simulation, which triggers a period of readaptation. This can also appear in the form of a long-term behavioral periodicity (see Figure 16).

## 5 Summary and Discussion

We have presented an integrated system that can explore and learn the emergent behaviors of a neuro-body-environment system coupled through physical embodiment by applying a novel chaotic search method. The whole system is treated as a single high-dimensional dynamical system using intrinsic chaotic dynamics as a driving force for the exploration of its own emergent patterns. The search process is completely deterministic and is able to selectively entrain the system orbit to one of the patterns by imposing goal directedness toward a desired behavior. Adaptive calibration of incoming sensor signals was established by using homeostatic sensory regulation. By adjusting the waveforms of input signals to be close to those of the neural activities, the synchronicity between the neural and physical system was enhanced, and the neural system was able to cope with an arbitrary robotic system. The regulation in the input system resulted in
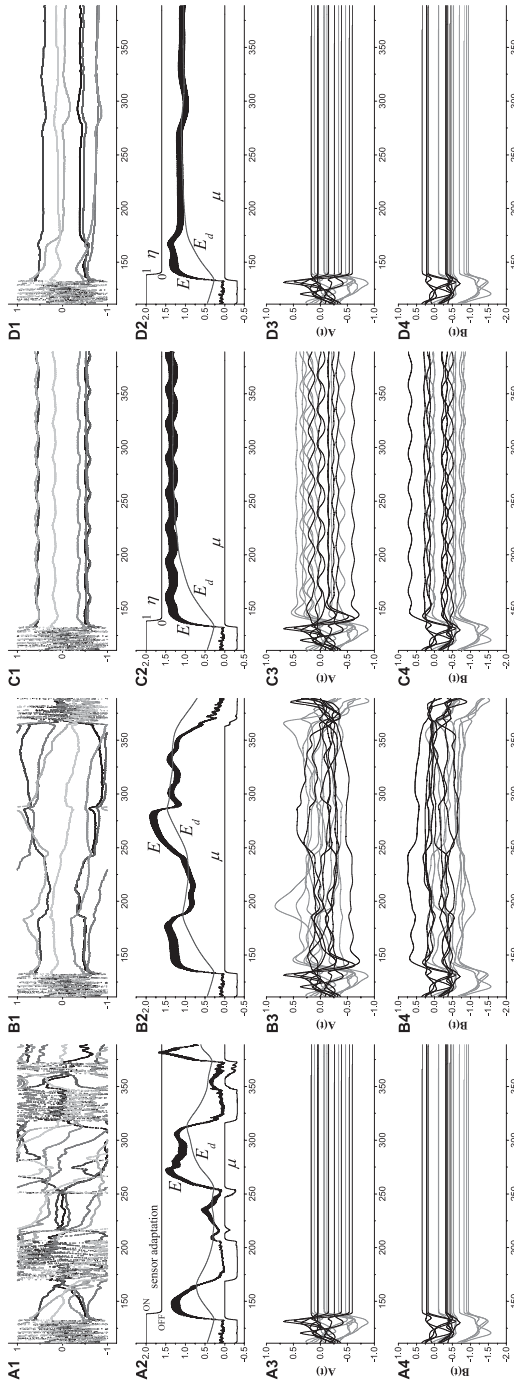
Figure 15: Four experiments (A–D) of quadruped exploration with different system settings starting from the same initial condition. (A1–D1) depict the phase differences between the limbs 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, and 1-8 (using different shading from black to light gray). The system behaves differently after the discovery of the first pattern ($t \approx 140$). The black and gray lines in A3–D3 and A4–D4 represent the parameters ($A(t)$, $B(t)$) of the sensor functions $H_l(s_l)$ and $H_r(s_r)$, respectively. (A1–A4) No sensor adaptation or oscillator learning was enabled. The sensor adaptation was turned off at the time of the first stabilization, and their values were maintained throughout the running period. The patterns degenerate quickly, and the system is hard to stabilize. (B1–B4) Sensor adaptation was turned on. The initial pattern changes much more slowly, and its performance stays at higher values for a long time until the system reenters the search phase following performance loss. (C1–C4) The pattern was captured by oscillator learning and completely stabilized. (D1–D4) Same as panel C, but the sensory adaptation was turned off as in panel A. The initially captured pattern was slightly shifted after fixing sensor parameters and showed slow degeneration over a long time.
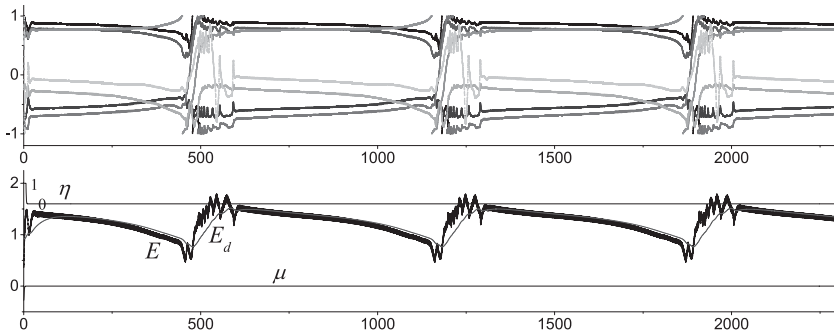
Figure 16: Long-term periodicity of quadruped behavior observed in a small number of cases. The coordinated pattern (straight walking) slowly shifts over a period of hundreds of cycles, followed by a short catastrophic transition (circling motion by asymmetric gait) and then reentrained to the same walking pattern.

the diversification of output behaviors in which the same neurosensory coordination could be achieved by different limb movements, accomplishing multiscale exploration. The discovered rhythmic pattern is memorized and sustained by wiring initially disconnected oscillators using an adaptive synchronization method. The oscillator learning process was naturally merged with the exploration system by using the emergent pattern as a supervising signal and could capture both stable and transient locomotor patterns in real time.

The overall process from the perspective of creating a new behavior can be briefly sketched as follows. The mutual entrainment between the neural and physical systems initially creates a phase space that contains several stable and transient patterns. If the current entrained state is not satisfactory, the system bifurcates to a chaotic state in order to escape from that state and restabilizes when a desired pattern appears. However, the phase space of the restabilized system differs from the previous one because some of the system parameters (sensor parameters) have also been changed by the chaotic drive. If we define the onset of stabilization (at the time $\mu$ becomes 0) as the time of returning, whenever the state orbit returns to the target space, it never experiences exactly the same phase space as before. This process is what we call multiscale exploration, and its eventual behavior after the onset of stabilization varies over different physical embodiments. The final dynamics of sensor adaptation after returning involves each parameter being locked around a particular value (potentially different for each parameter) with small oscillations. This diversity of parameter convergence can be regarded as the neutral stability of the system since different motor movements can cause the same sensory input. For the case of the 2D swimmer, which has a small number of strong basins of attraction, the

sensor parameters tend to converge to one of the previous distributions, although their precise values may differ. The neutrality in the convergence of sensor parameters has a wider range in the case of the quadruped; hence more diverse stabilized behaviors are exhibited. Even in the case where the sensor parameters eventually converge to the same set of distributions, the intermediate trajectories before convergence can take various routes, which can be captured by oscillator learning, resulting in the creation of a new behavior. Therefore, this process differs from a simple action selection mechanism where predetermined stable patterns are selected by a chaotic jump. Rather, it creates various streams of transient patterns by driving both the state orbit and the system parameters using chaotic dynamics.

Although our system has demonstrated a good degree of generality and an ability to automatically adapt to unknown bodies and environments, further analysis is necessary in order to determine the optimum values of fixed parameters used in the search process. For example, the timescales of slow dynamics such as evaluation ($\tau_E$), goal seeking ($\tau_d$), sensor adaptation ($\tau_h$), and feedback bifurcation ($\tau_\mu$) affect the search dynamics. Preliminary results of investigating the effect of different timescales revealed that the ratio between the timescales for evaluation, goal seeking, and feedback bifurcation determines the balance between the memorizing and forgetting of patterns during the search process (Aida & Davis, 1994), implying there might be an optimal ratio that allows the system to stay in the chaotic regime for an optimal duration (just enough to be uncorrelated with the previously visited pattern), enabling fast search with a very small probability of being trapped in a bad state for a long time. The timescale of the sensor adaptation can influence the landscape of phase space as well as the neutrality of convergence. A test using the 2D swimmer showed that when $\tau_h$ was decreased by $1/2$, a new stable pattern appeared where the two arms moved with large amplitudes, whereas the movements of the other two were small and irregular. However, too fast a timescale caused large fluctuations in parameters, which disturbed stabilization or diminished the diversity of behaviors synchronized with the fast state dynamics. Another factor that influences the system is the amount of bandwidth of the information flow between neural elements mediated by physical embodiment, which is determined by the design of body-environment interactions. In the case of the 4-fin swimmer presented here, the functional coupling strength between motor units varies with body mass. Increased body mass will result in an increased moment of inertia, which causes less transmission of the hydraulic force from one leg to the others, and vice versa. A similar effect will be caused by decreasing the density of the surrounding fluid or by increasing fin stiffness.

As Kuniyoshi and Sangawa (2006) have stated, completely decoupled CPGs are an extreme model, which might deviate from biological reality. However, some biological studies point out the evidence for functional decoupling of the neural system during certain phases or behaviors. It

has been hypothesized that decoupling of locomotor CPGs (as in our system) serves as a potential mechanism for the evolution of novel behaviors (Dubbeldam, 2001). Motion analysis of *Siren lacertina*, an eel-like amphibian (Azizi & Horton, 2004) has found strong evidence that the axial and appendicular CPGs are decoupled during aquatic walking (a pattern somewhere between aquatic and terrestrial locomotion), which supports the hypothesis that the decoupling of CPGs has led to the evolution of this novel behavior. In a broader perspective, Rosslenbroich (2009) pointed out that the locomotor neural processes of more evolved vertebrates are uncoupled from one another so that these parts can act in more differentiated and partly independent ways, which may contribute to the increase in organismic autonomy necessary for evolutionary innovation.

These emergent patterns may be refined and selected at the supraspinal level by reward-based reinforcement, which is thought to be one of the primary functions of the basal ganglia (BG) (Redgrave, Prescott, & Gurney, 1999; Schultz, 2006; Chakravarthy, Joseph, & Bapi, 2010). Recent modeling studies on BG (Sridharan, Prashanth, & Chakravarthy, 2006; Magdoom et al., 2011) hypothesize that the indirect striato-pallidal pathway through the subthalamic nucleus subserves exploratory behavior for goal-directed learning, gated by the dopamine signal from the substantia nigra which serves as the global learning signal for reward prediction. We hypothesize that goal-directed chaotic exploration may possibly take a role in such mechanisms in connection with self-organized behaviors. In this context, it might be possible to use our system to draw some implications about optimal parameters in relation to metalearning and neuromodulation centered around the BG (Doya, 2002).

Recent work has demonstrated the efficacy of morphological change within the context of locomotion behaviors created through an evolutionary search process (Bongard, 2011). An interesting area of future research will be to investigate whether the advantages of growth and development demonstrated in that work carry over to our method.

The system has also been successfully tested on other kinds of robots using identical neural controllers with the quadruped, further demonstrating its generality (see the movies presented in the URL in footnote 1). Although the final movement patterns produced by our work are never poor, they are not always perfectly optimized. Future work will explore the use of slightly more complex evaluation signals in this context. Also, we intend to incorporate adaptation to external perturbations, such as dealing with nonstationary environments. This might be achieved by using another adaptive system on top of the learned locomotion controller, or it may well be possible to develop such behavior within a slightly extended version of the current system. More intelligent and complex locomotion behavior could be achieved by using conventional learning methods or fuzzy control in conjunction with the concepts encapsulated in our system. The novel neuro-robotic system presented in the letter has been shown to

be general and effective. The seamless interaction between the exploration and learning processes results in a system that can be thought of as continually self-monitoring in order to maintain an appropriate level of motor function. As well as being an effective means of developing robotic controllers, the method has more general implications for truly autonomous artificial systems, which must maintain their integrity on several levels, including behavioral. The work demonstrates the possibility of the spontaneous emergence of meaningful behaviors in a continuous dynamical system framework, an approach that deviates from conventional learning algorithms making use of repeated trials.

## Appendix

**A.1 Oscillator Learning.** Let us consider $M$ oscillators that are fully connected to each other. We denote the state $j$ of the oscillator $i$ as $x_i^j$ and write a compact expression for equations 3.1 and 3.2 with the coupling term $F$,

$$\tau \frac{dx_i^j}{dt} = f^j(\mathbf{x}_i, t) + F_i^j, \quad j = 1, 2, \tag{A.1}$$

where $\mathbf{x}_i = (x_i^1, x_i^2)^T$ is the state vector. The sensory input term was regarded as part of the oscillator dynamics to promote sensory influence in the global coordination of the learned oscillator network. Assuming that the oscillators produce sinusoidal waveforms, the phase-locked solution of the state vector of oscillator $i$ and those of the other oscillators can be expressed as a linear relationship,

$$\mathbf{x}_i = \sum_{k=1}^{M} \mathbf{P}_{ik}\mathbf{x}_k, \quad i \neq k, \tag{A.2}$$

where $\mathbf{P}_{ik}$ is the $2 \times 2$ phase-lock matrix for the oscillators $i$ and $k$. Suppose we already have a certain phase relationship between $\mathbf{x}_i$ and other oscillators during the exploration process; then we can drive $\mathbf{x}_i$ in order to satisfy the equality in equation A.2 by using a simple error feedback to the oscillator using the gradient of an objective function $E_i$ such that

$$E_i = \frac{1}{2}\left\| \mathbf{x}_i - \sum_{k=1}^{M} \mathbf{P}_{ik}\mathbf{x}_k \right\|^2 = \frac{1}{2}\sum_{j=1}^{2}\left\{ x_i^j - \sum_{k=1}^{M}\sum_{l=1}^{2} p_{ik}^{jl}x_k^l \right\}^2, \tag{A.3}$$

$$F_i^j = -g\frac{\partial E_i}{\partial x_i^j} = g\left( \sum_{k=1}^{M}\sum_{l=1}^{2} p_{ik}^{jl}x_k^l - x_i^j \right). \tag{A.4}$$

$p_{ik}^{jl}$ represents the $(j, l)$th element of the matrix, and $g$ is a feedback gain, which should be set small enough so that the ongoing oscillation is not distorted. Thus, we can rewrite equation A.1 by neglecting the small decay term $gx_i^j$ in equation A.4 as

$$\tau \frac{dx_i^j}{dt} = f^j(\mathbf{x}_i, t) + g \sum_{k=1}^{M} \sum_{l=1}^{2} p_{ik}^{jl} x_k^l \quad i \neq k. \tag{A.5}$$

We can see that the feedback term represents the coupling term from other oscillators in that $g p_{ik}^{jl}$ is the coupling connection strength from $x_k^l$ to $x_i^j$. The coupling matrix $\mathbf{P}_{ik}$ can be obtained using the same gradient descent learning with regard to $p_{ik}^{jl}$. In order to eliminate any bias effect, the deviation of signal from its temporal average ($\tau_E \frac{d\bar{x}}{dt} = -\bar{x} + x$) was used for learning,

$$\frac{dp_{ik}^{jl}}{dt} = -\gamma \frac{\partial E_i}{\partial p_{ik}^{jl}} = \gamma \left\{ (x_i^j - \overline{x_i^j}) - \sum_{r=1}^{M} \sum_{s=1}^{2} p_{ir}^{js}(x_r^s - \overline{x_r^s}) \right\} (x_k^l - \overline{x_k^l}), \tag{A.6}$$

where $\gamma$ is the adaptive learning rate.

**A.2 Robot Simulation.** The robot simulations were implemented using open dynamics engine (Smith, 1998). The CPG and other differential equations were integrated using the Runge-Kutta (4th order) method with a step size of 0.0025 sec (the ODE simulation used the same step size). All code was written in C++.

A torsional muscle is activated by the $\alpha$ motoneuron (see Figure 2) and produces forces that cause movements of the body. We employed a simple yet biologically relevant actuation model proposed by Ekeberg (1993) and Wadden and Ekeberg (1998), where the motoneuron output linearly controls a muscle spring constant. The torque exerted on a joint by a pair of symmetric muscles is

$$T = \alpha(\sigma_l - \sigma_r) - \beta(\sigma_l + \sigma_r + \gamma)\theta - \delta\dot{\theta}, \tag{A.7}$$

where $\sigma_l$ and $\sigma_r$ are the output signals from the two motoneurons and $\theta$ is the joint angle. $\alpha$ is the muscle activation gain, $\beta$ is the stiffness gain, $\gamma$ is the tonic stiffness, and $\delta$ is the damping coefficient. Although they can be chosen arbitrarily from a wide range of values, we set those parameters for the muscles at rest to behave close to a critically damped system (see the values in Table 2). The above arrangement enables the neural signal not only to control the muscle torque but also to control the muscle tone.

Each motoneuronal output $\sigma$ is contributed to by CPG output ($x$) and a simplified muscle stretch reflex ($s_m$) according to the following canonical

Table 2: Robot Simulation Parameters.

| 4-Fin Swimmer | | Quadruped | |
| --- | --- | --- | --- |
| Torso dimension (m) | 0.2×0.2×0.2 | Torso dimension (m) | R: 0.05, L:0.9 |
| Arm Dimension (m) | 0.075×0.075×0.15 | Leg dimension (m) | R: 0.05, L:0.3 |
| Torso weight (Kg) | 1.6 | Torso weight (Kg) | 7.6 |
| Arm weight (Kg) | 0.34 (×4) | Leg weight (Kg) | 1.44 (×8) |
| Joint range (rad) | ±0.25π | Joint range (rad) | Upper: ±0.15π |
| Fin dimension (m) | 0.2×0.2 | | Lower: ±0.1π |
| Fin weight (Kg) | 0.375 | Friction coefficient | 1.0 |
| Fin stiffness (Nm) | 0.1 | Muscle parameters | |
| Fin damping (Nms) | 0.045 | $\alpha$ (Nm) | 7.935 |
| Fluid density (Kg/m$^3$) | 1000.0 | $\beta$ (Nm) | 1.684 |
| Muscle parameters | | $\gamma$ (Nm) | 20.0 |
| $\alpha$ (Nm) | 1.076 | $\delta$ (Nms) | 1.156 |
| $\beta$ (Nm) | 0.108 | | |
| $\gamma$ (Nm) | 20.0 | | |
| $\delta$ (Nms) | 0.152 | | |

formulas based on the literature (Prochazka, 1999; Yakovenko, Gritsenko, & Prochazka, 2004):

$$\sigma_{l,r} = \tanh(x \mp k_m s_m), \tag{A.8}$$

$$s_m = \theta/\Phi + R(\dot{\theta}/\Phi), \tag{A.9}$$

$$R(v) = \operatorname{sgn}(v)\sqrt{|v|}. \tag{A.10}$$

$k_m = 0.1$ is a constant, and $\Phi$ is a denominator that normalizes the angle and the angular velocity of torsional muscle by the unit of its resting angle, and it is set to the maximum available joint angle ($\Phi = \frac{2\alpha}{\beta\gamma}$) by assuming that the angle of the torsional muscle is stretched twice as much as its resting angle when the joint is at its neutral position. Although several types of proprioceptive feedback mechanisms, including groups Ia, Ib, II, and cutaneous afferents, operate on the spinal reflex system and their collective interaction accounts for the regulation of ongoing locomotor activities (Grillner & Wallen, 1985; Hiebert & Pearson, 1999; Pearson, Ekeberg, & Büschges, 2006; Rossignol, Dubuc, & Gossard, 2006), it would be sufficient to support the mechanical stability of muscles using a minimal model for a basic reflex loop since the group Ia pathway is the most sensitive of all. From the viewpoint of the global system, even the muscular-motoneuronal reflex loop can be broadly considered as part of the intact anatomical properties that may vary across the different robotic designs, which should be covered by the exploration process.

*A.2.1 4-Fin Swimmer.* The swimmer was constructed using a 3D rigid body simulator, but it was constrained to move only on the $x$-$y$ plane, so that it effectively undergoes 2D dynamics. Each fin of the 2D swimmer was modeled as a nonlinear damped torsional spring, which is subject to simulated hydrodynamics (Shim & Kim, 2006), and its bending angle ($\phi$) was fed to the corresponding motor unit. The fin angle implements the stretch receptor at each side of the fin, so the afferent inputs $s$ in equations 3.1 to 3.4 were defined as $s_l = \phi$ and $s_r = -\phi$. By assuming that a fin sensor reflects the output difference of the oscillator pair in the corresponding motor unit (i.e., $s_{l,r} = f(x_{r,l} - x_{l,r})$), we use the following slightly reformulated CPG equations for the 4-fin swimmer.

$$\tau \dot{x}_{l,r} = c \left( x_{l,r} - \frac{x_{l,r}^3}{3} - y_{l,r} + z_{1,2} \right) + \delta H_{l,r}(s_{l,r}) + F_{l,r}^x, \tag{A.11}$$

$$\tau \dot{y}_{l,r} = \frac{1}{c}(x_{l,r} - by_{l,r} + a) + \varepsilon (H_{l,r}(s_{l,r}) + x_{l,r}) + F_{l,r}^y. \tag{A.12}$$

Thus, the reference neural signal for sensory adaptation in equation 3.6 should also be changed to $n_l = -n_r = x_r - x_l$. The time constant and the maximum bifurcation parameter used were $\tau = 0.8$ and $\mu_c = 0.35$. All other parameters are as defined in equations 3.1 to 3.4.

*A.2.2 Quadruped.* The quadruped body was configured as bilaterally symmetric, and the Coulomb friction model with a coefficient of 1.0 was used. It has 8 motor units (16 oscillators), which is as many as the number of degrees of freedom. The sprawl posture of the legs and the proper setting of joint ranges ensure that the robot will not overturn during exploration. The CPG equations described in equations 3.1 to 3.4 are used unaltered for the quadruped. Because there are two separate sensor signals from each muscle, in this case the sensor design matches the forms of the CPG equations. $\tau = 0.4$ and $\mu_c = 0.32$ were used for the controller. The stretching force (pulling torque) that a muscle experiences (the collective effect not only of the antagonistic muscle but also the whole-body motion) was chosen as the appropriate sensory signal, which contains well-blended neural-body-environmental information. At any given time instance, if a torsional muscle is stretched by $\phi$ from its resting angle, the pulling torque that the muscle receives can be expressed using the muscle activation level ($\sigma$, which linearly controls the muscle spring constant) and the stretched angle. This relationship can be written in a processed form using the joint angle $\theta$ as

$$s_{l,r} = (\sigma_{l,r} + 1) \left( 1 \mp \frac{\theta}{\Phi} \right) - 2. \tag{A.13}$$

The equation was centered around zero to have a range [−2, 2] and fed to the relevant SAM. Neurobiological studies on the spinal or decerebrate cat (Conway, Hultborn, & Kiehn, 1987; Pearson, Ramirez, & Jiang, 1992; Pearson & Collins, 1993; Pearson, 2008) suggest that a major component of sensory signals from the ankle extensor muscle receptors, which influence the central rhythm-generating network, is related to the force in the muscles, which primarily arises from the Golgi tendon organs (group Ib afferent), while the group Ia afferent from muscle spindles has a lesser effect.

**A.3 Simulation Parameters.** The detailed parameters for robots and physical simulation are described in Table 2.

## Acknowledgments

## References

Aida, T., & Davis, P. (1994). Oscillation mode selection using bifurcation of chaotic mode transitions in a nonlinear ring resonator. *IEEE Transactions on Quantum Electronics*, *30*(12), 2986–2997.

Asai, Y., Nomura, T., Abe, K., & Sato, S. (2003). Classification of dynamics of a model of motor coordination and comparison with Parkinson's disease data. *Biosystems*, *71*, 11–21.

Asai, Y., Nomura, T., Sato, S., Tamaki, A., Matsuo, Y., Mizukura, I., et al. (2003). A coupled oscillator model of disordered interlimb coordination in patients with Parkinson's disease. *Biological Cybernetics*, *88*, 152–162.

Azizi, E., & Horton, J. M. (2004). Patterns of axial and appendicular movements during aquatic walking in the salamander *Siren lacertina*. *Zoology*, *107*, 111–120.

Barlas, Y., & Yasarcan, H. (2006). Goal setting, evaluation, learning and revision: A dynamic modeling approach. *Evaluation and Program Planning*, *29*(1), 79–87.

Bongard, J. (2011). Morphological change in machines accelerates the evolution of robust behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(4), 1234–1239.

Bongard, J., Zykov, V., & Lipson, H. (2006). Resilient machines through continuous selfmodeling. *Science, 314*, 1118–1121.

Buchli, J., Righetti, L., & Ijspeert, A. J. (2006). Engineering entrainment and adaptation in limit cycle systems. *Biological Cybernetics*, *95*(6), 645–664.

Chakravarthy, V. S., Joseph, D., & Bapi, R. S. (2010). What do the basal ganglia do? A modeling perspective. *Biological Cybernetics*, *103*(3), 237–253.

Cohen, A. H., Rossignol, S., & Grillner, S. (1988). *Neural control of rhythmic movements in vertebrates*. New York: Wiley.

Conway, B. A., Hultborn, H., & Kiehn, O. (1987). Proprioceptive input resets central locomotor rhythm in the spinal cat. *Experimental Brain Research*, *68*, 643–656.

Davis, P. (1990). Application of optical chaos to temporal pattern search in a nonlinear optical resonator. *Japanese Journal of Applied Physics*, *29*, L1238–L1240.

Doya, K. (2002). Metalearning and neuromodulation. *Neural Networks*, *15*, 495–506.

Doya, K., & Yoshizawa, S. (1992). Adaptive synchronization of neural and physical oscillators. In J. E. Moody, S. J. Hanson, & R. P. Lippman (Eds.), *Advances in neural information processing systems, 4* (pp. 109–116). San Francisco: Morgan Kaufmann.

Dubbeldam, J. L. (2001). Evolution of playlike behaviour and the uncoupling of neural locomotor mechanisms. *Netherlands Journal of Zoology*, *51*, 335–345.

Ekeberg, O. (1993). A combined neuronal and mechanical model of fish swimming. *Biological Cybernetics*, *69*, 363–374.

Feudel, U., Grebogi, C., & Yorke, J. A. (1996). Map with more than 100 coexisting low-period periodic attractors. *Physical Review E*, *54*, 71–81.

Floreano, D., Husbands, P., & Nolfi, S. (2008). Evolutionary robotics. In B. Siciliano & O. Khatib (Eds.), *Springer handbook of robotics* (pp. 1423–1451). New York: Springer.

Freeman, W. J., & Viana Di Prisco, G. (1986). EEG spatial pattern differences with discriminated odors manifest chaotic and limit cycle attractors in olfactory bulb of rabbits. In G. Palm & A. Aertsen (Eds.), *Brain theory* (pp. 97–119). New York: Springer-Verlag.

Gallagher, J. C., Beer, R. D., Espenschied, K. S., & Quinn, R. D. (1996). Application of evolved locomotion controllers to a hexapod robot. *Robotics and Autonomous Systems*, *19*, 95–103.

Grillner, S., & Wallen, P. (1985). Central pattern generators for locomotion control, with special reference to vertebrates. *Annual Review of Neuroscience*, *8*, 233–261.

Guevara, M. R., Glass, L., Mackey, M. C., & Shrier, A. (1983). Chaos in neurobiology. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*, 790–798.

Hiebert, G. W., & Pearson, K. G. (1999). Contribution of sensory feedback to the generation of extensor activity during walking in the decerebrate cat. *Journal of Neurophysiology*, *81*(2), 758–770.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A*, *454*, 903–995.

Ijspeert, A. J. (2001). A connectionist central pattern generator for the aquatic and terrestrial gaits of a simulated salamander. *Biological Cybernetics*, *84*(5), 331–348.

Ijspeert, A. J., Crespi, A., Ryczko, D., & Cabelguen, J. M. (2007). From swimming to walking with a salamander robot driven by a spinal cord model. *Science*, *315*(5817), 1416–1420.

Itoh, Y., Taki, K., Kato, S., & Itoh, H. (2004). A stochastic optimization method of CPG-based motion control for humanoid locomotion. In *Proceedings of the IEEE Conference on Robotics, Automation and Mechatronics* (pp. 347–351). Piscataway, NJ: IEEE.

Kamimura, A., Kurokawa, H., Toshida, E., Tomita, K., Murata, S., & Kokaji, S. (2003). Automatic locomotion pattern generation for modular robots. In *Proceedings of the IEEE International Conference on Robotics and Automation* (pp. 714–720). Piscataway, NJ: IEEE.

Kaneko, K. (1994). Relevance of dynamic clustering to biological networks. *Physica D*, *75*(3), 55–73.

Kaneko, K. (2003). Chaotic itinerancy. *Chaos*, *13*(3), 926–936.

Kelso, J.A.S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Kimura, H., Akiyama, S., & Sakurama, K. (1999). Realization of dynamic walking and running of the quadruped using neural oscillator. *Autonomous Robots*, *7*, 247–258.

Kinjo, K., Nabeshima, C., Sangawa, S., & Kuniyoshi, Y. (2008). A neural model for exploration and learning of embodied movement patterns. *Journal of Robotics and Mechatronics*, *20*(3), 358–366.

Korn, H., & Faure, P. (2003). Is there chaos in the brain? II. Experimental evidence and related models. *Comptes Rendus Biologies*, *326*, 787–840.

Kuniyoshi, Y., & Sangawa, S. (2006). Early motor development from partially ordered neural-body dynamics: Experiments with a cortico-spinal-musculoskeletal model. *Biological Cybernetics*, *95*, 589–605.

Kuniyoshi, Y., & Suzuki, S. (2004). Dynamic emergence and adaptation of behavior through embodiment as coupled chaotic field. In *Proceedings of IEEE International Conference on Intelligent Robots and Systems* (pp. 2042–2049). Piscataway, NJ: IEEE.

Magdoom, K. N., Subramanian, D., Chakravarthy, V. S., Ravindran, B., Amari, S., & Meenakshisundaram, N. (2011). Modeling basal ganglia for understanding Parkinsonian reaching movements. *Neural Computation*, *23*(2), 477–516.

Matsubara, T., Morimoto, J., Nakanishi, J., Sato, M., & Doya, K. (2006). Learning CPG-based biped locomotion with a policy gradient method. *Robotics and Autonomous Systems*, *54*, 911–920.

Morihiro, K., Isokawa, T., Matsui, N., & Nishimura, H. (2005). Reinforcement learning by chaotic exploration generator in target capturing task. *Lecture Notes in Computer Science*, 3681 (pp. 1248–1254). Berlin: Springer-Verlag.

Mpitsos, G. J., Burton, R. M., Creech, H. C., & Soinila, S. O. (1988). Evidence for chaos in spike trains of neurons that generate rhythmic motor patterns. *Brain Research Bulletin*, *21*, 529–538.

Nakamura, Y., Mori, T., Sato, M., & Ishii, S. (2007). Reinforcement learning for a biped robot based on a CPG-actor-critic method. *Neural Networks*, *20*(6), 723–735.

Nara, S., & Davis, P. (1992). Chaotic wandering and search in a cycle-memory neural network. *Progress of Theoretical Physics*, *88*(5), 845–855.

Nomura, T., Sato, S., Doi, S., Segundo, J. P., & Stiber, M. D. (1993). A Bonhoeffer–van der Pol oscillator model of locked and non-locked behaviors of living pacemaker neurons. *Biological Cybernetics*, *69*, 429–437.

Ohgane, K., Ei, S. I., & Mahara, H. (2009). Neuron phase shift adaptive to time delay in locomotor control. *Applied Mathematical Modelling*, *33*(2), 797–811.

Ohgi, S., Morita, S., Loo, K., & Mizuike, C. (2008). Time series analysis of spontaneous upper-extremity movements of premature infants with brain injuries. *Physical Therapy*, *88*(9), 1022–1033.

Ott, E., Sauer, T., & Yorke, J. A. (1994). *Coping with chaos: Analysis of chaotic data and the exploitation of chaotic systems*. New York: Wiley.

Parker, T. S., & Chua, L. O. (1989). *Practical numerical algorithms for chaotic systems*. New York: Springer-Verlag.

Pearson, K. G. (2008). Role of sensory feedback in the control of stance duration in walking cats. *Brain Research Review*, *3139*, 1–26.

Pearson, K. G., & Collins, D. F. (1993). Reversal of the influence of group Ib afferents from plantaris on activity in medial gastrocnemius muscle during locomotor activity. *Journal of Neurophysiology*, *70*(11), 1009–1017.

Pearson, K. G., Ekeberg, O., & Büschges, A. (2006). Assessing sensory function in locomotor systems using neuro-mechanical simulations. *Trends in Neurosciences*, *29*(11), 625–631.

Pearson, K. G., Ramirez, J. M., & Jiang, W. (1992). Entrainment of the locomotor rhythm by group Ib afferents from ankle extensor muscles in spinal cats. *Experimental Brain Research*, *90*, 557–566.

Pfeifer, R., & Bongard, J. (2007). *How the body shapes the way we think: A new view of intelligence*. Cambridge, MA: Bradford Books.

Pitti, A., Lungarella, M., & Kuniyoshi, Y. (2005). Quantification of emergent behaviors induced by feedback resonance of chaos. *Advances in Natural Computation*, *15*(3), 199–213.

Pitti, A., Niiyama, R., & Kuniyoshi, Y. (2010). Creating and modulating rhythms by controlling the physics of the body. *Autonomous Robots*, *28*(3), 317–329.

Prochazka, A. (1999). Quantifying proprioception. *Progress in Brain Research*, *123*, 133–142.

Raftery, A., Cusumano, J., & Sternad, D. (2008). Chaotic frequency scaling in a coupled oscillator model for free rhythmic actions. *Neural Computation*, *20*(1), 205–226.

Rapp, P., Zimmerman, I., Albano, A., Deguzman, G., & Greenbaun, N. (1985). Dynamics of spontaneous neural activity in the simian motor cortex: The dimension of chaotic neurons. *Physics Letters A*, *110*(6), 335–338.

Redgrave, P., Prescott, T., & Gurney, K. N. (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience*, *89*, 1009–1023.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–69). New York: Appleton-Century-Crofts.

Riley, M., & Turvey, M. (2002). Variability and determinism in motor behaviour. *Journal of Motor Behavior*, *34*(2), 99–125.

Rossignol, S., Dubuc, R., & Gossard, J. P. (2006). Dynamic sensorimotor interactions in locomotion. *Physiological Review*, *86*, 89–154.

Rosslenbroich, B. (2009). The theory of increasing autonomy in evolution: A proposal for understanding macroevolutionary innovations. *Biology and Philosophy*, *24*(5), 623–644.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Physiology*, *57*, 87–115.

Shim, Y. S., & Husbands, P. (2010). Chaotic search of emergent locomotion patterns for a bodily coupled robotic system. In *Proceedings of 12th International Conference on the Synthesis and Simulation of Living Systems (ALIFE XII)* (pp. 757–764). Cambridge, MA: MIT Press.

Shim, Y. S., & Kim, C. H. (2006). Evolving physically simulated flying creatures for efficient cruising. *Artificial Life*, *12*(4), 561–591.

Skarda, C., & Freeman, W. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, *10*, 161–195.

Smith, R. (1998). *Open dynamics engine (ODE)*. Available online at http://ode.org/.

Sproewitz, A., Moeckel, R., Maye, J., & Ijspeert, A. J. (2008). Learning to move in modular robots using central pattern generators and online optimization. *International Journal of Robotics Research*, *27*(3), 423–443.

Sridharan, D., Prashanth, P. S., & Chakravarthy, V. S. (2006). The role of the basal ganglia in exploration in a neural model based on reinforcement learning. *International Journal of Neural Systems*, *16*, 111–124.

Stein, P. S. G., Grillner, S., Selverston, A., & Stuart, D. G. (Eds.). (1997). *Neurons, networks and motor behavior*. Cambridge, MA: MIT Press.

Steingrube, S., Timme, M., Worgötter, F., & Manoonpong, P. (2010). Self-organized adaptation of a simple neural circuit enables complex robot behaviour. *Nature Physics*, *6*, 224–230.

Terman, D., & Rubin, J. (2007). Neuronal dynamics and the basal ganglia. *SIAM News*, *4*(2). Available online at http://siam.tekdevelopment.com/old–issues/2007/march–2007/.

Turrigiano, G. G. (2008). The self-tuning neuron: Synaptic scaling of excitatory synapses. *Cell*, *135*, 422–435.

Turrigiano, G. G., & Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, *5*, 97–107.

Vadivasova, T. E., Sosnovtseva, O. V., Balanov, A. G., & Astakhov, V. V. (1999). Phase multistability of synchronous chaotic oscillations. *Discrete Dynamics in Nature and Society*, *4*, 231–243.

Wadden, T., & Ekeberg, O. (1998). A neuro-mechanical model of legged locomotion: Single leg control. *Biological Cybernetics*, *79*, 161–173.

Wheeler, M. (2005). *Reconstructing the cognitive world: The next step*. Cambridge, MA: MIT Press.

Wright, J., & Liley, D. (1996). Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. *Behavioral and Brain Sciences*, *19*, 285–320.

Yakovenko, S., Gritsenko, V., & Prochazka, A. (2004). Contribution of stretch reflexes to locomotor control: A modeling study. *Biological Cybernetics*, *90*(2), 146–155.

Zhang, C. K., & Shao, H. H. (2001). A hybrid strategy: Real-coded genetic algorithm and chaotic search. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics 4* (pp. 2361–2364). Piscataway, NJ: IEEE.