Monostable Controllers for Adaptive Behaviour

Christopher L. Buckley[1], Peter Fine[2], Seth Bullock[1], and Ezequiel Di Paolo[2]

[1] School of Electronics and Computer Science, University of Southampton, UK
[2] School of Informatics, University of Sussex, UK
clb05r@ecs.soton.ac.uk, p.a.fine@sussex.ac.uk,
sgb@ecs.soton.ac.uk, ezequiel@sussex.ac.uk

Abstract. Recent artificial neural networks for machine learning have exploited transient dynamics around globally stable attractors, inspired by the properties of cortical microcolumns. Here we explore whether similarly constrained neural network controllers can be exploited for embodied, situated adaptive behaviour. We demonstrate that it is possible to evolve globally stable neurocontrollers containing a single basin of attraction, which nevertheless sustain multiple modes of behaviour. This is achieved by exploiting interaction between environmental input and transient dynamics. We present results that suggest that this globally stable regime may constitute an evolvable and dynamically rich subset of recurrent neural network configurations, especially in larger networks. We discuss the issue of scalability and the possibility that there may be alternative adaptive behaviour tasks that are more 'attractor hungry'.

Keywords: Global stability, echo state networks, evolvability.

1 Introduction

Certain regions of the cortex are organised into neural microcolumns. It has been suggested that the computational power of these cortical microcolumns stems from their transient dynamics rather than their attractor structure [1]. This could be the result of weak coupling between the neurons, both in terms of their connectivity and weight strength, which can confer global stability on their dynamics [2]; consider that ripples on the surface of a liquid are only a temporary echo of a stone dropped into a it. Any "computation" undertaken by such a system may only be achieved by exploiting interaction between its environmental input and its transient dynamics within a single basin of attraction. Recently, the echo state and liquid state approaches have employed artificial neural networks that share this constrained dynamics [1,3]. They have been shown to perform well on a range of machine learning tasks [4]. Despite this, and notwithstanding the biological heritage of these artificial neural networks (ANNs), it is not immediately obvious that they can serve as effective control systems for adaptive behaviour since, in general, such controllers must cope with an agent that is embodied and situated in a changing environment that demands different modes of behaviour in different circumstances.

Here we explore whether continuous-time recurrent neural networks (CTRNNs) constrained to exhibit transient dynamics around a globally asymptotically stable fixed

point attractor are capable of successfully completing a task that demands different behavioural modes. First, we briefly recap the role of attractor dynamics in the use and understanding of ANNs. Subsequently, we introduce a state-hungry task and evolve a solution that employs bistability. Section 3 presents results from evolutionary studies in which systems constrained to exhibit a single attractor are demonstrated to solve the task. After making a preliminary investigation of the difference between the two kinds of solution, we go on to consider the scalability of this type of constrained neural network. We conclude by discussing the implications of these results.

1.1   Mechanisms of State Retention in Neural Networks

Much of the early work on artificial neural networks focused on feedforward architectures [5]. Here, information is processed along a unidirectional pipeline mapping sensory input onto motor output. This process is necessarily atemporal and does not involve internal state. Stateful networks only really arrived in the 1980s, when Hopfield popularised recurrent neural networks (RNNs). Hopfield's (1982) RNNs were constrained such that each was guaranteed to exhibit a number of fixed point attractors. In such networks, an input is a static or slowly time-varying signal and the system is allowed to converge to an equilibrium. The particular attractor that is achieved is interpreted as a recognition or recall event. The success of Hopfield style systems resulted in the attractor becoming the dominating neurological metaphor for memory and state within biological organisms.

More recent work in neuroscience and adaptive behaviour research have reconsidered organisms as dynamical systems, and foregrounded the role of time [7]. This has led to the frequent observation that, far from settling into stable attractor states, neural systems often spend most of the time exhibiting *transient* dynamics, often far from equilibrium. Systems that tend to settle into static states when examined in isolation (either in a petri dish or as computational models) may in fact be far from equilibrium when coupled to bodies and environments. However, while one need not subscribe to the idea that attractors within the agent's control system dynamics are 'representational', intuitively, it may still be attractive to interpret the presence of distinct attractors as enabling for tasks that demand distinct behavioural modes.

Here we probe this intuitive correspondence between behavioural modes and distinct attractors by comparing control systems that are constrained to exhibit a single attractor with those that are capable of exhibiting multiple fixed-point and/or cyclic attractors. Before defining the task, agent architecture and evolutionary scheme that will be employed for the remainder of the paper, we introduce and formalise some constraints on global asymptotic stability.

1.2   Single Fixed-Point Attractors and the Echo State Property

An echo state machine, comprising a large neural network "reservoir" and a simple feedforward network "readout", must satisfy the echo state property, typically achieved by enforcing global asymptotic fixed-point stability (GAS). Informally, the echo state property demands that the dynamics of the recurrent neural network (the reservoir) will "wash out" all information from the initial conditions after some sufficient period of

time [3]. One formulation of this condition demands that the distance between any two trajectories in the system is a monotonically decreasing function of time, i.e., the system exhibits a single limit set [3]. Jaeger provides a *sufficient* condition for GAS, however, it is difficult to check efficiently. Furthermore, it is believed to be overly restrictive since many systems which do not meet this strict condition may exhibit GAS. In practice, a simpler but only *necessary* condition can be derived for systems with an equilibrium coincident with the centre of each node's transfer function (i.e. centre-crossing networks [8]). While not sufficient for GAS, this condition has been shown to confer the echo state property in the majority of echo state applications [3].

In this paper we modify this condition so that it can be used with a CTRNN equation given by:

$$\tau_i \dot{y}_i = -y_i + \sum_j \omega_{ji} \tanh(y_j + \theta_j) + I_i \tag{1}$$

Here y_i represents the activation at the ith neuron, ω_{ji} is the weight of the connection from neuron j to neuron i, θ_i is the bias value at the ith neuron, τ_i defines the rate of leakage or decay of the activation, and $\tanh$ is the transfer function. While this equation is not identical to either Jaeger or Beer's formulations, all such networks fall in the larger class of Cohen and Grossberg networks [9] to which the following stability results will apply.

As stated above, we first require that the network is in the centre-crossing configuration. While this is not trivial to impose in general [8], for Equation (1) it can be achieved by setting all of the system's biases to zero ensuring that there is an equilibrium at the zero state, i.e., $\mathbf{y}^* = \bar{\mathbf{0}}$, where $\mathbf{y}^*$ is a vector describing the position of an equilibrium point. We then construct a criterion for local stability by linearising the system around this point. Given that the slope of the $\tanh$ function is equal to unity at its centre, the linear dynamics is completely described by a Jacobian at the equilibrium point given by:

$$J = \begin{pmatrix} \frac{\omega_{11}-1}{\tau_1} & \cdots & \frac{\omega_{1N}}{\tau_1} \\ \vdots & & \vdots \\ \frac{\omega_{N1}}{\tau_N} & \cdots & \frac{\omega_{NN}-1}{\tau_N} \end{pmatrix} \tag{2}$$

This system will be locally stable around the zero state if all real parts of the eigenvalues of the Jacobian are negative, otherwise it will be unstable [10]. Local stability is a necessary but not sufficient precondition for global stability of the full, nonlinear centre-crossing system [11]. For a full discussion of the condtions for local and global stability in CTRNNs see [10].

2   Methods

In order to experimentally investigate the role of attractor structure in an embodied, be-having agent, we needed to construct a task that demands state and encourages multiple modes of behaviour. We carried out experiments using a simulated agent required to perform phototaxis towards a sequence of lights using a single light sensor. The task was made more challenging by alternating the location of the light sensor intermittently

between the front and back of the agent's body throughout its lifetime. Since the agent cannot reverse, the nature of the agent's phototactic task is fundamentally altered by this unsignalled change to its sensor. In previous work on this task evolved solutions employed bistable controllers [12].

The experiment consists of a circular agent with radius equal to unity located in an infinite 2-D plane. (Distance and time are measured in arbitrary units.) At the start of each trial, a single light source is presented, located at a random distance, d from the agent, drawn from a uniform random distribution over the range [10, 15], in any random direction. After a variable time period drawn uniformly at random from the range [40, 60], the trial finishes, and a new trial commences with a new randomly re-positioned light. Less frequently, the sensor is switched from the front to the back of the agent (or vice-versa) at the start of a trial.

The agent is controlled by a neural network receiving input from the light sensor and driving two motors, which differentially steer the robot with their output (in range [0, 1]). The light sensor accepts incoming light so long as it is not occluded by the agent's body, and provides a value, I, in [0, 1] varying inversely with the distance between the sensor and the light source up to a maximum of 150% of the largest possible initial distance between the agent and the light.

The agent is controlled by a network of continuous time recurrent neurons governed by Equation (1). Time constants (τ_i) were scaled as e^{x_i} with x_i drawn from the uniform distribution [0, 5], and weights w_{ij} and biases θ_i drawn from the uniform distribution [-10, 10]. The sensor value is scaled by a sensor weight s uniform in [-10, 10], and is made available to the first neuron, only. The outputs of each of the last two neurons were used to generate left and right motor speeds, respectively.

Motor neuron outputs are first scaled by a motor weight, ω_r, and translated by a motor bias, θ_r, before being squashed and rescaled in the range [0, 1] to prevent reversing. Thus, the speed of the right motor, r, was derived from the output of the right motor neuron, y_r, as $r = \frac{1}{2}[1 + \tanh(\omega_r y_r + \theta_r)]$. The network (and other simulation variables) are integrated with an Euler time-step of 0.1 during optimisation of the agent's controller, and 0.01 during analysis (to ensure stability).

Network parameters were optimised using a genetic algorithm employing pairwise tournament selection and a population of 50, for up to 6000 generations. A losing genotype was replaced with a copy of the winner subject to parameter mutation via Gaussian perturbation (zero mean, variance scaled between 0.02 and 0.05 as the previous generation's elite genotype fitness varies between 0.4 and 0.8). Fitness was calculated as the normalised average distance of the agent from the light during the last 25 time units of each trial. The first trial after the sensor was switched does not contribute towards fitness, allowing for a possible adaptation phase to occur without punishing the agent.

Evolution progressed according to a shaping scheme. When the best agent of each of the 15 prior generations had attained a fitness greater than 0.8, the phase was advanced. During phase one, the sensor was solely located on the front of the agent, so standard phototaxis was all that was required. During its lifetime, the agent was subject to eight sets of six sequential light presentation trials, with its internal state reset between each set. The average of these eight sets was taken as the overall fitness score. Phase two consisted of the phase one presentations followed by an additional eight sets of six trials

with the sensor located on the back of the agent. The overall fitness score was the mean of all 16 sets. A third phase proceeded as per phase two, but with double the number of trials per set (i.e., 12). During each set of trials the sensor was switched to the opposite side of the agent at the start of a random trial number drawn uniformly from the range [4, 8]. In the final phase, three changes of sensor location took place at random intervals within each of 16 sets, each comprising 16 trials. As such, the successful completion of phase four demands that an agent must be able to cope with repeated alternation between front-mounted and rear-mounted sensors.

3   Results

Before exploring the behaviour of different recurrent networks on the task described above, we ran experiments with feedforward neural networks that are unable to exploit internal state. These networks comprised nodes governed by Equation (1). However, all recurrent connections were removed, all time constants were set to unity ($\tau_i = 1$), and each network's Euler integration step was also set to unity. Of 70 runs, not a single evolved controller was able to progress beyond phototaxis (the first phase of the shaping scheme). Based on these results, we conclude that it is difficult, or perhaps impossible for a reactive control system to solve the full behavioural task for the agent/environment combination explored here.

Unconstrained CTRNNs. We wish to determine whether the two behavioural modes that were exploited in previously reported work [12] result from bistability in the autonomous dynamics of such controllers, and whether this is a necessary property of evolved CTRNN solutions. To this end, we examined successfully evolved four-node CTRNNs and compared them with four-node CTRNNs that are biased towards having global stability and hence are less likely to express bistable dynamics.

It proved somewhat difficult to evolve controllers in this scenario. Of the 50 evolutionary runs evolving standard CTRNNs, only three agents were produced that were able to successfully complete the final phase of the task. In order to investigate whether the intrinsic dynamics of the successful control systems were bistable, we first considered their autonomous dynamics, i.e., in the absence of input. We tested for *autonomous global stability* by allowing each network to relax from 50 different random initial conditions ($y_i \in [-10, 10]$). This analysis revealed that the networks were not globally stable, revealing the presence of two fixed point attractors within the intrinsic dynamics of all three of the successful control systems.

The fact that more than one fixed-point attractor is present in a network's dynamics does not necessarily imply that the agent's internal state spends time in more than one basin of attraction during its behaving lifetime, i.e., a network might exhibit *lifetime global stability* in the absence of autonomous global stability. We tested for this possibility by running trials of each successful agent, and removing any sensory input at various stages during these trials. For all of the successful solutions, the agent's dynamics settled to one attractor when its light sensor was front-mounted, and the other attractor when it was rear-mounted.

Stable CTRNNs. We have described how, when permitted to explore the full CTRNN parameter space, successful solutions exhibited bistable lifetime dynamics. In order to determine whether it was possible to satisfy this task in a network that possessed only a single attractor in its autonomous dynamics, the stability constraint developed above (Section 1.2) was applied during a further 50 evolutionary runs. Recall: this is only a *necessary* condition for stability, and therefore can only bias the population towards globally stable fixed point (monostable) solutions.

The stability constraint was implemented by initialising all weights $\omega_{ij} \in [-0.3, 0.3]$, and rejecting any of these random networks that did not satisfy the constraint. While evolution was permitted to explore a full CTRNN weight range ($\omega_{ij} \in [-10, 10]$), any mutations that generated a network which failed to satisfy the stability constraint were rejected, and a new offspring was attempted.

From 50 runs, two evolved working solutions. In order to confirm that these solutions exhibited autonomous and/or lifetime global stability, we repeated the tests described above for unconstrained networks. For both solutions, the network's autonomous internal state always settled to same fixed point from every initial condition tested. Furthermore, we ensured that they exhibited autonomous global stability in the presence of constant input (over the range experienced by the agent during its lifetime). Straightforwardly, this carried over to stability in their lifetime dynamics, also.

3.1   Contrasting Solutions

How did the bistable and globally stable solutions differ from one another? Figure 1 depicts a projection of the internal dynamics of two successful controllers undergoing a series of trials during which the sensor's position is switched. The bistable solution (on the left) transitions between two basins of attraction, exhibiting two modes corresponding to the two types of phototactic behaviour demanded by the task. By contrast, while the dynamics of the stable system also exhibit two distinct quasi-oscillatory modes,
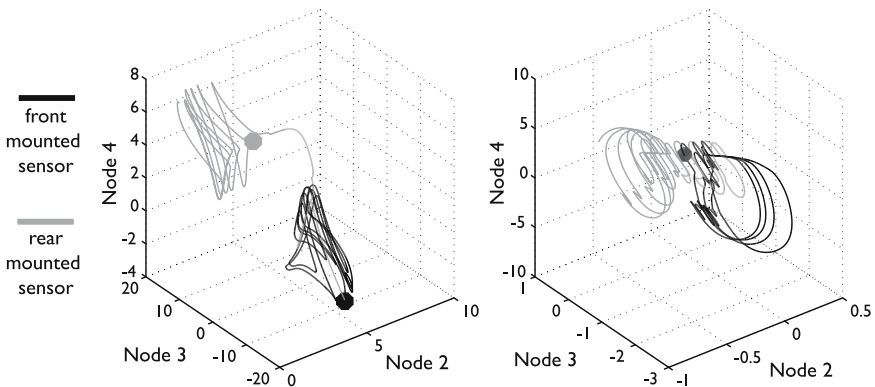

Fig. 1. Phase space plots of a typical trajectory during several consecutive trials of the best evolved unconstrained (left) and constrained (right) networks. Half way through, the sensor switches from the front to back. The discs indicate the attractor locations (in the absence of input); one (the origin) in the constrained case, and two in the bistable case.

these trajectories all occur within the same basin of attraction. In both cases, the transition from one basin to another is prompted by "pathological" sensory input, which occurs soon after the sensor location is changed. In the former case, this is responsible for moving the controller's internal state across a separatrix in the system's dynamics. In the latter case, rather than using environmental input to transition across a separatrix, the monostable controller uses it to shuttle between different regions of the same basin of attraction. Whereas a monostable system must rely solely upon this environmental coupling to maintain at least one behavioural mode, a bistable solution may exploit its *autonomous* dynamics to maintain two distinct modes of behaviour. However, accounting for the importance of the intrinsic dynamics in the complete brain/body/environment system is problematic. We will return to this issue in Section 4.

Figure 2 presents time series for two runs of the same monostable controller, one of which undergoes sensor relocation at the beginning of the second trial. In this trial the trajectories proceed identically until the light is first detected, at which point they begin to diverge. The unchanged agent's behaviour proceeds as before, with the positive slope of the bursts of sensory input indicating that the agent is approaching the light. However the neural trajectories diverge as the sensory input of the agent whose sensor has been switched at the start of this trial diminishes, due to the now maladaptive behaviour. By the third trial a different and adaptive pattern of behaviour is achieved and maintained.

3.2   Scalability and Stability

The four-node networks that we have considered so far are clearly much smaller than typical liquid state machines and echo state networks, and even smaller than the cortical
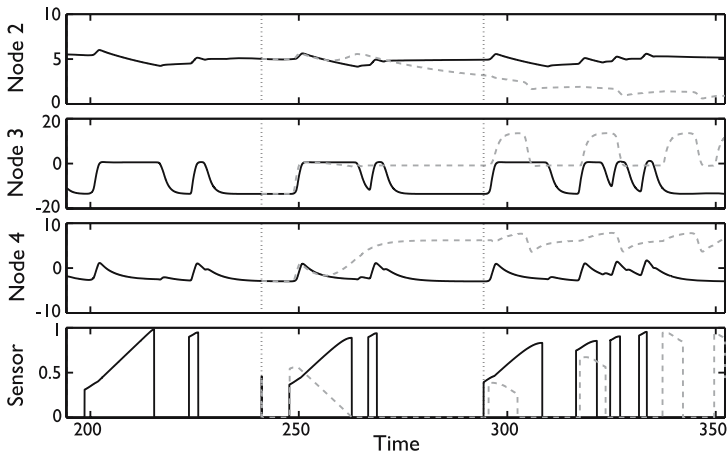

Fig. 2. Time series for three of the four nodes plus sensory input, during three consecutive trials (separated by dotted lines). The solid line depicts front-mounted sensor behaviour in three consecutive trials. The dashed line depicts rear-mounted sensor behaviour in trials two and three. At the outset of trial two, in both cases the light is positioned such that it is initially occluded from the sensor by the agent's body.

networks that inspired them. Like other kernel methods, the high dimensionality of these networks is core to their ability to use transient dynamics for pattern recognition and machine learning tasks. By contrast, high-dimensional CTRNNs are not typically employed for adaptive behaviour tasks. This stems partly from the computational demands of simulating large networks and partly from the analytic challenge that must be overcome in understanding their behaviour.

In addition, as network size increases, significant portions of the search space exhibit saturating dynamics that may be unhelpful for the production of interesting behaviour [13]. One possible method for encouraging interesting generic dynamics in large networks is to bias evolutionary search toward interesting regions of parameter space. Specifically, recall that in order to obtain the echo state property in the networks considered here, we have placed them in the centre-crossing configuration. In general, such networks have been shown to oscillate readily, making them an appropriate substrate for the evolution of, e.g., pattern generators [8]. Furthermore, it has been suggested that networks in the centre-crossing configuration will exhibit rich dynamics because their nodes interact at the most sensitive parts of their transfer functions. However, this oscillatory behaviour, which becomes more pronounced in large CTRNNs, can be disruptive, interfering with effective signal transduction. One possible solution is to further constrain them to exhibit global asymptotically stable fixed point behaviour, as we have done here. Such a constraint mitigates against destructive reverberative oscillation while retaining sensitivity to input. This has been shown to encourage effective signal propagation in large networks and has been conjectured to constitute a computationally rich subregion of CTRNN space [14].

To explore this we perform a preliminary study of the scalability of constrained versus unconstrained networks by repeating the evolutionary experiments reported above for networks comprising ten nodes. We also evolved unconstrained networks with all bias values set to zero in order to distinguish the contribution to performance of the stability constraint from that of the centre-crossing property. The results of 50 evolutionary runs per network type are reported in table 1.

Table 1. Success and stability rates for three classes of 10-node CTRNN

(50 runs)	Unconstrained	Constrained	Centre crossing
(i) % Successful Runs	28%	60%	44%
(ii) % of (i) that were Lifetime Stable	50%	90%	55%
(iii) % of (ii) that were Globally Stable	43%	89%	8%

Increasing the number internodes in the networks increases the number of successful evolutionary runs for each of three network classes. However, both the constrained and unconstrained centre-crossing systems produced more successful solutions than the unconstrained CTRNNs, with the constrained networks producing the greatest number overall. Furthermore, the majority of solutions even for unconstrained CTRNNS employed a single basin of attraction.

4 Discussion

One of the initial demands that we placed on our agent is the requirement for a control system that maintains some kind of internal state. It is important to be clear that we are not talking about any form of representation within the agents, but merely the ability to retain information over temporally extended periods. The construction of a task that *demands* state to be held specifically *within* a control system is problematic. For example, doubt has been cast on how much can be understood about the cognitive limitations of an agent's behaviour from the limitations on its internal dynamics. It has been shown that even purely reactive systems with no internal state are capable of behaviour that can be interpreted as non-reactive [15]. However, while many tasks could theoretically be satisfied with a reactive controller in interaction with a rich and dynamic environment, in practice, both biological and artificial control systems frequently exploit internal state. Consequently, while it may be hard to specify tasks that can *never* be satisfied with a purely reactive controller, it is likely that many tasks could be described as being at least 'state hungry', possibly because the agents involved do not have the necessary privileged access to their environment.

Adaptive behaviour research typically discusses the dynamics of internal state in terms of transients around the attractors of a system (even if the equilibirum associated with an attractor is never reached). In this paper we have made an attempt to understand the dynamics of transients in the absence of complex attractor structure. To this end we have examined a task that demands multiple behavioural modes. Given the bistability exhibited by evolved CTRNN solutions, it might be tempting to equate distinct attractors with distinct behavioural modes. However, we were subsequently able to evolve networks able to satisfy the same task with only a single fixed point attractor. Closer inspection reveals that both kinds of solution exhibit two distinct bundles of transients corresponding to the two behavioural modes. Given that the agent has recourse to some kind of environmental interaction with which to separate these two bundles, it need not rely on a seperatrix to differentiate these behaviours in its *autonomous* dynamics.

While monostable solutions to the task were evolved, for small CTRNNs they did not arise readily under a conventional evolutionary robotics methodology. Indeed, in order to obtain such solutions we had to explicitly encourage stable controllers. However, in larger networks (see Section 3.2) solutions that utilised a single basin of attraction evolved readily, even when the networks were unconstrained. One possible implication is that the utility of monostable versus multi-stable controllers may be sensitive to network size.

Lastly, might there be classes of behavioural task that *cannot* be satisfied without the presence of multiple attractors? In the same way that one can conceive of a task as being 'state hungry', might a particular subset of tasks be 'attractor hungry'? Consider the challenge posed by multiple time-scales of adaptive behaviour. While a bistable controller can retain certain state information indefinitely, by virtue of relaxing to one of many stable states, a monostable controller must rely on environmental stimuli on an appropriate timescale. Conversely, opting to solve a task by employing a minimal number of basins of attraction may also afford certain advantages in terms of evolvability, tunability and generalisability. This is supported by the fact that, for the task considered here, unconstrained CTRNNs often evolve to exploit a single basin of attraction,

and that constraining CTRNNs in this way improves their chances of evolving a successful controller. However, further studies of the interaction between network scaling and evolvability are necessary before we can confirm that in order for larger networks to achieve their full computational potential, they will benefit from mechanisms that constrain them into a stable centre-crossing configuration.

References

1. Maass, W., Natschläger, T., Markram, H.: Real-time computing without stable states: A new framework for neural computation based on perturbations. Computation 14, 2531–2560 (2002)
2. Hirsch, M.W.: Convergent activation dynamics in continous time networks. Neural Networks 2, 331–349 (1989)
3. Jaeger, H.: The echo state approach to analysing and training recurrent neural networks. GMD-Report 148, German National Research Institute for Computer Science (2001)
4. Jaeger, H., Maas, W.: Special issue on echo state networks and liquid state machines. Neural Networks 20, 287–289 (2007)
5. Rumelhart, D.E., McClelland, J.L.: Parallel Distributed Processing: Explorations in the Microstructure of Cognition. MIT Press, Cambridge (1986)
6. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. In: Proceedings of the National Academy of Sciences, USA, vol. 79, pp. 2554–2558 (1982)
7. Beer, R.D.: The dynamics of active categorical perception in an evolved model agent (with commentary and response). Adaptive Behavior 4(11), 209–243 (2003)
8. Mathayomchan, B., Beer, R.D.: Center-crossing recurrent neural networks for the evolution of rhythmic behavior. Neural Computation 14, 2043–2051 (2002)
9. Cohen, M., Grossberg, S.: Absolute stability of global pattern formation and parrallel memory storage by competetive neural networks. IEEE Trans on Syst, Man and Cyber 13, 815–826 (1983)
10. Strogatz, S.H.: Nonlinear Dynamics & Chaos. Addison-Wesley, Reading (1994)
11. Jaeger, H., Lukosevicius, M., Popovici, D., Siewart, U.: Optimization and application of echo state networks with leaky-integrator neurons. Neural Networks 20, 335–352 (2007)
12. Fine, P., Di Paolo, E., Izquierdo, E.: Adapting to your body. In: Costa, F.A.e., Rocha, L.M., Costa, E., Harvey, I., Coutinho, A. (eds.) Ninth European Conference on Artificial Life, pp. 203–212. Springer, Heidelberg (2007)
13. Williams, H., Noble, J.: Homeostatic plasticity improves signal propagation in continuous time recurrent neural networks. Biosystems 87(2-3), 252–259 (2007)
14. Buckley, C.L., Bullock, S.: Sensitivity and stability:a signal propagation sweet spot in a sheet of recurrent centre crossing neurons. Submitted to Biosystems (2007)
15. Izquierdo-Torres, E., Di Paolo, E.: Is an embodied system ever purely reactive? In: Capcarrere, M., Freitas, A., Bentley, P.J., Johnson, C.G., Timmis, J. (eds.) Eighth European Conference on Artificial Life, pp. 252–261. Springer, Heidelberg (2005)