Is an Embodied System Ever Purely Reactive?

Eduardo Izquierdo-Torres and Ezequiel Di Paolo

Centre for Computational Neuroscience and Robotics,
Deparment of Informatics, University of Sussex
{e.j.izquierdo, ezequiel}@sussex.ac.uk

Abstract. This paper explores the performance of a simple model agent using a reactive controller in situations where, from an external perspective, a solution that relies on internal states would seem to be required. In a visually-guided orientation task with sensory inversion and an object discrimination task a study of the instantaneous response properties and time-extended dynamics explain the non-reactive performance. The results question common intuitions about the capabilities of reactive controllers and highlight the significance of the agent's recent history of interactions with its environment in generating behaviour. This work reinforces the idea that embodied behaviour exhibits properties that cannot be deduced directly from those of the controller by itself.

1 Introduction

Is it possible to deduce the cognitive limitations of an embodied agent from the limitations of its internal dynamics? In particular, is an agent controlled by a reactive system able to perform only reactive behaviours? Questions like these compel us to look carefully at the meaning of now commonly used terms such as embodiment and situatedness, often discussed in the abstract, and try to unravel their implications for concrete systems.

A way of addressing the most specific of these questions is to build agents controlled by reactive systems and evaluate their performance in situations that require non-reactive responses. By a *reactive controller* we understand a system whose outputs are at each moment only determined by its current inputs. In order to make the problem non-trivial we need to define reactive behaviour in terms of the properties of the *task* and not the controller. For the purpose of this work we adopt a definition of reactive behaviour based on the classification introduced by Clark and Thornton [4] as the performance of a type-1 task, i.e., a task that requires the agent to exploit regularities which are directly apparent in the current input data. In robotics, obstacle avoidance is typically a type-1 task. In contrast, type-2 tasks require the exploitation of regularities which are 'hidden', or whose statistical visibility depends on some systematic recoding of the data. Accordingly, we will treat performance of a type-2 task as a form of *non-reactive behaviour*. Online learning is typically a type-2 task.

In this paper, evolutionary algorithms are used to design neurocontrollers for the behaviour of model agents which are then analysed dynamically. The goal is

M. Capcarrere et al. (Eds.): ECAL 2005, LNAI 3630, pp. 252–261, 2005.

to explore the role of embeddedness in space and time in enabling non-reactive performance in systems that can only respond reactively. In particular, we investigate the relation between instantaneous response properties and time-extended performance in orientation tasks, and the time-dependence of responsiveness and 'decision making' in shape discrimination. In both cases, embodied agents exhibit properties that cannot be deduced directly from their reactive controllers. The dynamical analysis of these agents allows us to draw some general inferences about the danger of making *a priori* assumptions about the required properties of internal control mechanisms for a given task.

2 Background

We may find classical answers to our opening questions in criticisms of behaviourism. For instance, in Dewey's critique of the reflex-arc concept in psychology [5] it becomes clear that action is ongoing and stimuli can only have an effect on the behaving agent because the agent is capable of selecting them actively by the nature of its perceptual systems but also by the nature of its actions. The same point is compellingly made by Merleau-Ponty:

> "The organism cannot properly be compared to a keyboard on which the external stimuli would play and in which their proper form would be delineated for the simple reason that the organism contributes to the constitution of that form ... it is the organism itself – according to the proper nature of its receptors, the threshold of its nerve centers and the movement of the organs – which chooses the stimuli in the physical world to which it will be sensitive." (in [9] p.13)

We find similar views in Varela's work (e.g. [13]), where the emphasis is on cognition as embodied action wherein the world and the perceiver mutually specify each other. This is closely related to von Uexküll's functional circles [8], i.e. the formation of a closed unit between the 'perceptual' and 'effector' worlds that enables an agent to generate its own *Umwelt*. In robotics, Pfeifer subscribes to a related view [11], showing the importance of thinking in terms of sensorimotor coordinations. In recent years, we have seen concrete examples of these ideas at work in the area of autonomous robotics. For instance, Nolfi [10] provides examples in object size and shape classification tasks using reactive controllers and Scheier et al. [12] make similar points by studying object-constancy and focus-of-attention problems using hand-coded physical robots as well as evolved simulated agents. Implications of the embodied view in the context of biological neural networks have been summarised in [3].

What we propose to do here is to focus on the opportunities that the 'neural' architecture, body and environment offer to the system's controller. We will show how and why an embodied system can perform non-reactive behaviour (type-2 tasks) even when only endowed with a purely reactive controller. The interesting lessons will be in the details of how the agents work because they uncover the hidden assumptions about the capabilities of embodied and situated systems, even when their internal controllers are very simple.

3 Methods

We propose to study the role of the agent's situatedness using a set-up similar to the one presented in [1,2] with slight variations and extensions on the architecture and tasks. This model has been chosen for two reasons: its simplicity and its potential for sufficiently interesting behaviours that could be called minimally cognitive. As a modification to this set-up, the controller's architecture is made purely reactive. Two tasks are studied: an approach/avoid task which is made type-2 by the fact that the sensor array may be inverted, and a discrimination task to evaluate the reactivity of the agent at different stages. These are described in the following sections.

The agent has a circular body with a diameter of 30. The agent's 'eye' consists of six rays at $\pm\pi/12$, $\pm\pi/24$ and $\pm\pi/36$ from the centre. An intersection between a ray and an object causes an input to be injected into the corresponding sensory node, with the magnitude of the injected input inversely proportional to the distance to the object along that ray with a maximum length of 220. When the rays are at their maximum length, no input is injected, while the maximum input is injected for rays of zero length.

The agent can move horizontally as objects fall from above (Figure 1A) with horizontal velocity proportional to the sum of the opposing forces produced by two motors. The behaviour of the agent is controlled by a network of continuous-time recurrent neural nodes of the form:

$$\tau_i \dot{y}_i = -y_i + \sum_{j=i}^{N} w_{ji} \sigma(g_j\,(y_j + \theta_j)) + I_i \tag{1}$$

where y is the activation of each node, τ is its time constant, w_{ji} is the strength of the connection from the j^{th} to the i^{th} node, θ is a bias term, g is a gain, $\sigma(x) = 1/(1 + e^{-x})$ is the standard logistic activation function, I represents a constant external input (e.g. from a sensor) and N is the number of nodes in the network. In simulation, node activations are calculated forward through time by straightforward time-slicing using Euler integration with a time-step of 0.1.

The network architecture is bilaterally symmetric in the connection weights, biases and time-constants (unless otherwise specified). The architecture consists of six ray sensory nodes projecting to five inter-nodes, which in turn project to two motor nodes controlling horizontal motion (Figure 1B). All the sensory nodes share the same time-constant, bias parameter and gain parameter, while the rest of the nodes have a gain of 1.

The controller is made reactive by changing the connection weights between the inter-nodes to 0, fixing the time-constants for all nodes to 1, and modifying the time-step of integration to 1. In this way the state of any node in the network is fully determined by the pattern of inputs and cannot depend on previous internal states resulting in a discrete-time feed-forward artificial neural network.

The parameters for the neural controllers are evolved using a microbial genetic algorithm [7]. These are encoded as a vector of real numbers over the range [0, 1] (47 parameters for recurrent controllers and 26 for reactive ones). Offspring

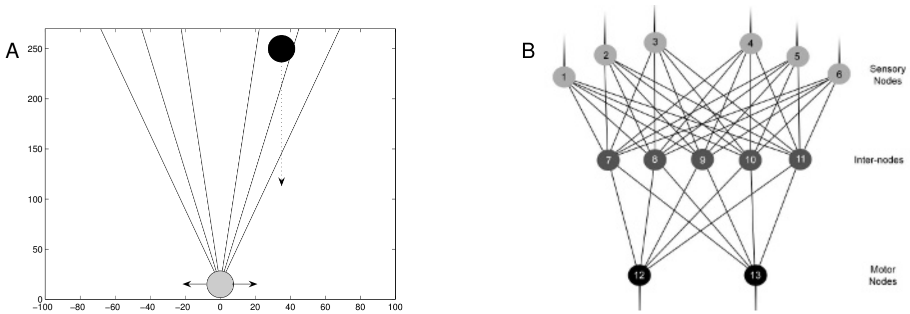

Fig. 1. Basic setup for the experiments. [A] The agent (gray circle) can move horizontally while objects (black circle) fall from above. The agent uses an array of 6 distal sensors (black lines). [B] The network architecture consists of six sensory nodes fully connected to five inter-nodes, which are in turn fully connected to two motor nodes.

of microbial tournaments are generated using recombination and vector mutation which consists of adding to the genotype a random displacement vector whose direction is uniformly distributed on the N-dimensional hypersphere and whose magnitude is a Gaussian random variable with 0 mean and variance 0.01. Population sizes of 100 and recombination rate of 0.9 are used. Genotypes are then mapped to network parameters using linear maps from $[0, 1]$ to $[-14, 14]$ for biases, $[-10, 10]$ for connection weights and $[1, 5]$ for the gain parameter while time-constants are exponentially mapped to $[e^0, e^4]$.

4 Orientation Experiments with Visual Inversion

In the first set of experiments, visually-guided agents are evolved to adjust their horizontal position so as to catch or avoid falling objects with normal and inverted vision. On inverting the visual field in the left/right direction an object that appears to the right of the agent will in fact be to its left. This task represents a type-2 problem, for it requires an agent to perform differently for the same stimuli depending on the context.

A simple evolutionary training regime is used. During an evolutionary evaluation 21 circular objects are dropped from the top of the environment straight down with an initial horizontal offset from the centre of the agent uniformly distributed in ± 50 and a fixed vertical velocity of -3. Following [6], this is repeated for objects of different diameter (ie. 26, 30 and 34). The whole process is then repeated using inverted vision, for a total of 126 trials in a fitness evaluation. At the start of each new trial node activations are initialised to zero.

The performance measure to be maximised is: $f = 1 - \sum_{i=1}^{N} d_i/N$, where N is the total number of trials and d_i is the horizontal distance between the centres of the object and the agent when their vertical separation goes to 0 on the ith trial (clipped to $max = 50$ and normalised).

Agents with a reactive controller that could orient to falling circles with normal and inverted vision turned out to be relatively easy to evolve. Over 20 evolutionary runs, the best evolved agent achieved a mean performance of 0.994 on the 126 evaluation trials after only 100 generations of evolution, with a mean performance of 0.992 using normal vision and 0.990 with inverted sensors on 10.000 randomly generated trials distributed with initial horizontal positions in $[-50, 50]$, and diameter and vertical velocity of the falling object between $[20, 40]$ and $[-4, -2]$, respectively.

Figure 2A shows the strategy used by the agent to catch falling circles with normal and inverted vision. Notice the opposed shading of the velocity fields in the two conditions. As the controller is reactive and symmetric, a stimulus produces instantaneous motor effects that are opposite in the case of normal and inverted vision, or put differently, a real and a virtual object in the same position produce exactly the same instantaneous effect. Yet the situated behaviour of the agent over time results in trajectories that catch both virtual and real objects.

In the normal condition, trajectories are attracted to the centre where the velocity field turns slightly divergent and then 'trapped' by the two bright regions of centring velocities which eventually converge on the object's horizontal position. In the inverted condition, central trajectories become convergent by the nature of the central field, and the rest of the trajectories initially move away from the centre only to be trapped in a different and wider convergent region, reaching the centre when the divergent fields no longer have the same effect. The evolved strategy involves taking advantage of the agent's multiple sensors and most successfully evolved agents relied on a very similar strategy.

Recurrent and time-based networks were evolved as well and analyses of the best evolved controller yielded the use of a similar strategy to that of the above analysed reactive network. Figure 2B shows the behaviour of the best evolved recurrent network over 20 evolutionary runs.

Agents with network architectures not constrained to be bilaterally symmetrical seemed to be relatively easier to evolve. The behaviour of the best evolved agent is shown in Figure 2C. The agent's strategy is to position itself sufficiently to one side of all falling objects, at which point real objects are seen with its right-sided sensors while virtual objects with its left set of sensors. The agent can then centre in on objects with opposite reactions according to which side they appear to be on. The result is a much simpler strategy for centring on both real and virtual objects.

A reactive agent needs to constantly engage with sensory stimuli in order to act which makes avoiding (as opposed to catching) falling objects with normal and inverted vision a counterintuitive task. Figure 2D shows the behaviour and dynamics of the best evolved reactive controller for such task. From the figure its strategy can be easily understood: under normal vision, the agent avoids objects that are far away and centres on objects that are relatively close. As a result, real objects get avoided as they start falling and disappear from the field of view early, while virtual objects are initially centred, reaching then the point were sufficiently closed objects get avoided.

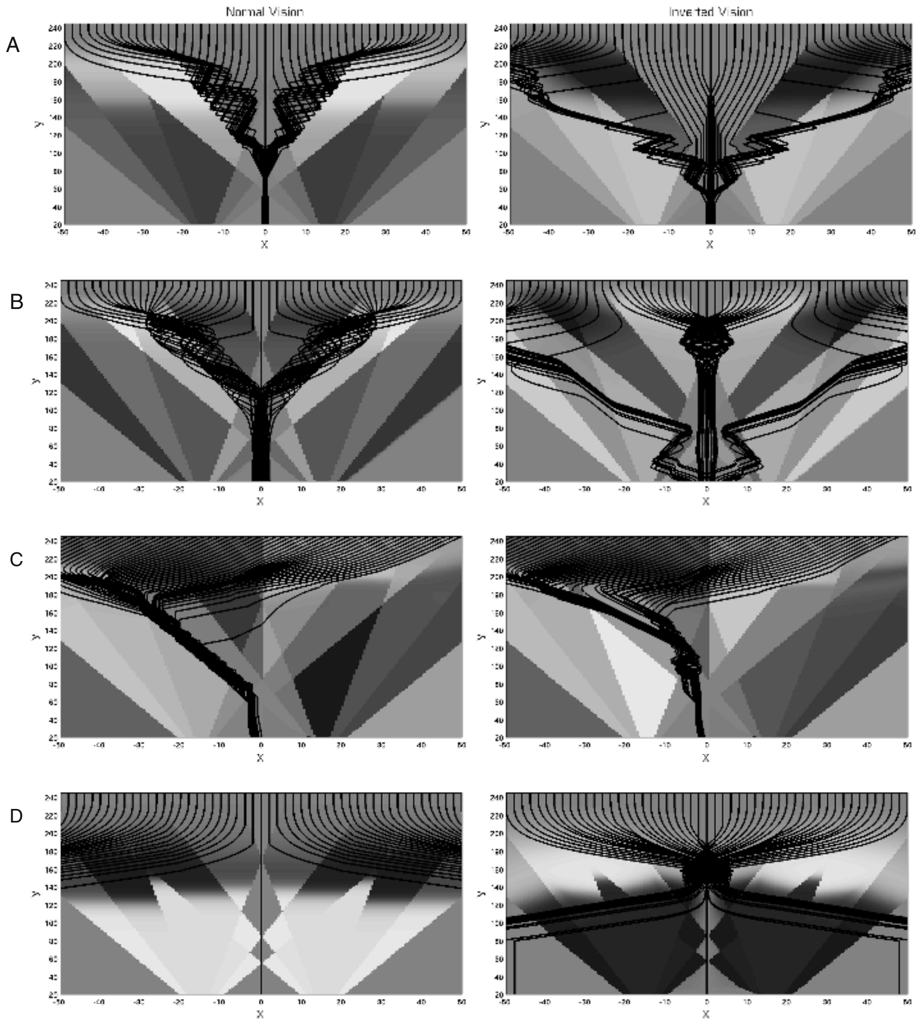

Fig. 2. Trajectories and steady-state horizontal velocity fields for normal (left) and inverted (right) vision for the best evolved agents for objects of diameter 30 and 51 different initial horizontal offsets for: [A] Circle centring task with a reactive symmetrical network; [B] Circle centring task with a recurrent symmetrical network; [C] Circle centring task with a reactive non-symmetric network; and [D] Circle avoidance task using a reactive symmetrical network. Trajectories are superimposed on differently shaded regions representing the long-term horizontal velocity that the agent would eventually adopt if an object were fixed at that location in its field of view at each point as a function of x the horizontal offset in relation to the agent and y the vertical distance of the object. Regions in which the agent is directed towards the object (centring regions) are bright, whereas those in which it is directed away (avoidance regions) are dark. The magnitude is represented by the intensity of the shade (mid-gray is no movement).

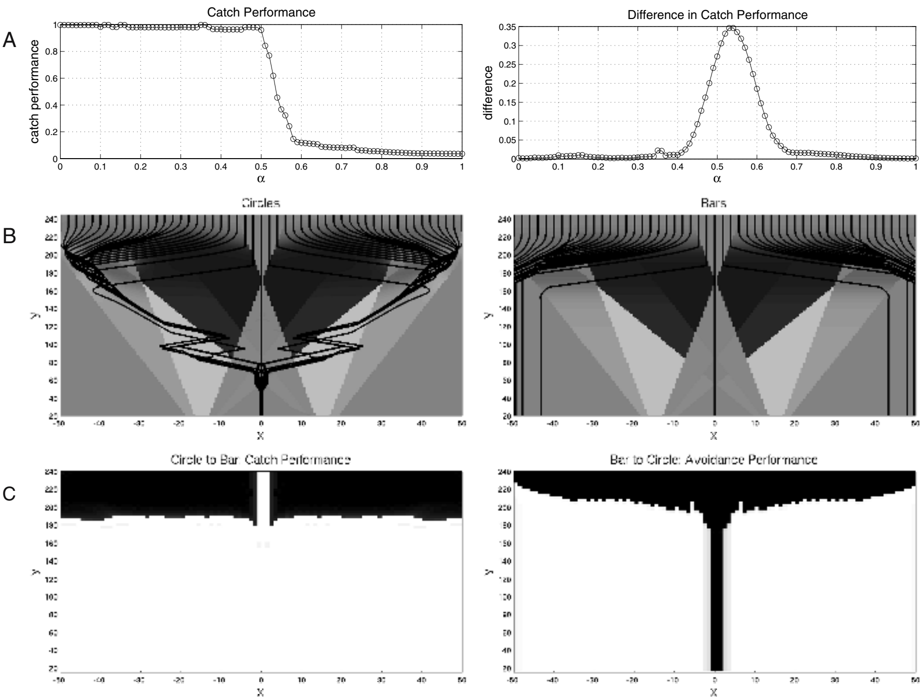

Fig. 3. [A] Demonstration of labelling and discrimination in the best evolved feedforward network. Average (left) and difference (right) in catch performance as a function of α. Each point represents the mean value for 101 trials at uniformly distributed horizontal offsets in ± 50. [B] Trajectories and steady-state horizontal velocity fields for the best evolved agent for circles (left) and bars (right). [C] Performance as a consequence from switching the object's identity from a circle into a bar (left) and viceversa (right) as a function of at different times on the final decision to catch or avoid in the best evolved agent. Each point represents the catch (left) and avoidance (right) performance as a function of initial horizontal offsets (x) and switching times (y). Bright means high performance. All figures are for objects of diameter 30 and vertical velocity -3.

5 Categorical Perception Experiments

In a second set of experiments, we explore agents that can discriminate between circular and horizontally oriented bars of different sizes using normal vision, catching the former and avoiding the latter in a similar task to the one explored in [2,6], in this case using a reactive controller. The evolutionary training regime used was similar to that used in the first set of experiments, with the only difference that half of the trials corresponded to circular falling objects and the other half to bar objects (as opposed to sensory inversion).

The performance measure to be maximised is: $f = p_i/N$, where N is the total number of trials and $p_i = 1 - d_i$ for a circular object and $p_i = d_i$ for bars, d_i is defined and normalised as above. Following [2], a class of parametrised hybrid object that continuously interpolates between the bar and the circle is defined as $R(\alpha) = \alpha R_b + (1 - \alpha)R_c$, where $R(\alpha)$ is the radius of the hybrid object, and R_b and R_c are the radii (in polar coordinates) of the bar and circle respectively.

Over 20 evolutionary runs, the best evolved agent achieved a mean performance of 0.970 on the 126 evaluation trials after 100 generations of evolution, with a mean performance of 0.915 on 10.000 randomly generated trials from a broader set (initial horizontal positions between $[-50, 50]$, diameter size of the falling object between $[20, 40]$ and vertical velocity between $[-4, -2]$).

Two major defining characteristics of categorical perception are labelling and discrimination. In order to demonstrate these, the mean catch or avoid performance was plotted as the shape of the object changed between a circle and a bar (by modifying the parameter α). Figure 3A depicts the average catch performance as a function of α, a sharp sigmoidal curve with a transition from catching to avoidance behaviour at about $\alpha = 0.55$ is observed. Accordingly, the average difference in catch performance for α values that differ by 0.1 as a function of α shows a bell shaped function.

How are we to explain the behaviour of the agent? What sort of regularities does it exploit from the environment? The behaviour and steady-state dynamics of this agent are shown in Figure 3B. The evolved strategy involves positioning all falling objects at a particular angle of view where the difference between the two objects is maximised. This can be appreciated from dominating dark regions in the middle-top field of view of the steady-state velocity. At the point where the object is positioned close to the border of the agent's field of view, circular objects fall onto a very thin bright region of centring behaviour. This is further explained from a closer look at what the agent 'sees' (figure not shown), a circle never stimulates less than 2 sensors, while the bar stimulates only 1 sensor at one point, and this makes it move out of the sensor range.

An interesting question in the context of this paper is: at what point during the trial does the agent commit itself to catching or avoiding an object? What is expected from a reactive agent is a strong dependence, throughout the trial, between the shape of the object and the 'decision' to catch or avoid. This is explored by switching the shape of the object during the trial and observing the behaviour. In the absence of an internal state to 'retain' a previously made decision, one expects the decision to depend mainly on the shape after the switch.

Figure 3C shows the performance of the agent when catching a circle or avoiding a bar as a function of the horizontal offset and the distance where the switch from circle to bar, or vice versa, is introduced. The results are contrary to our expectations. Although the agent seems to be completely uncommitted during the initial movements, after passing a more or less well defined 'decision line' it becomes largely committed to either catching or avoiding even if the shape is changed afterwards. The 'decision process' is very much a discrete event that occurs in the early stages of the trial.

The intuition goes wrong because it generalizes from the instantaneous effect of a pattern of stimuli on a reactive controller to the time-extended and situated behaviour of the agent. If, as explained above, discrimination is achieved by a particular correlation between object shape and angle of sensing chosen by the agent, and if after that event, independently of the decision made, the agent is already positioned in either a neutral or a centring velocity field, then any subsequent change of shape will be ignored. This is because behaviour does not depend on the objective shape of the stimulus but more precisely on the sensorimotor correlation between object and agent.

6 Discussion and Conclusions

This paper has demonstrated the evolution of embodied agents with reactive controllers for visually guided-orientation with sensory inversion and object discrimination. Although the tasks are interesting in themselves, the point of this paper is not to generate novel behaviour but to probe the intuitions concerning the capabilities of reactive controllers.

This work provides concrete examples showing how an embedded system is never purely reactive. From the example of shape discrimination, we show that the evolved agent will exploit state arising from its interaction with the environment and exhibit commitment to a decision. Agents modify their position with respect to other objects in the environment and, thus, partially determine the sensory patterns they will receive in the next time-step, thereby providing a concrete example of an agent creating the form of the stimulus by its manner of offering itself to actions from the outside, paraphrasing Merleau-Ponty.

For the visual inversion experiment the agent relies on following time-extended dynamics. As the state of the controller depends only on the pattern of inputs, the velocity fields for the normal and inverted conditions are point-by-point, opposed to each other. Which does not mean that the final state of the whole trajectory will be different in each case. This prompts an important conclusion: *the limitations of reactive controllers (or generally any given class of controllers) during extended periods of time in a situated system cannot be trivially deduced from the instantaneous, snapshot limitations of the same controllers.* Inversion of the sensory array produces an instantaneous reversal of velocities, and yet it results in a conservation, not a reversal, of the end-state of behaviour.

We illustrate some of the implications of reducing the assumptions about the necessary design of the agent's internal control mechanisms. In the visual inversion scenario, losing the symmetrical constraints allows the agent to *redefine* the problem into an easier one: catching objects that fall only to one side of it.

We do not deny the importance of an agent's internal dynamics in the generation of behaviour. It may, nevertheless, be the case that agents with internal dynamics exploit first the strategies available from its situatedness alone. In the visual inversion experiments agents with internal dynamics have the potential to solve the task in a variety of ways – for example, learning online which way around the sensors are wired up – and then acting accordingly. It is likely, however, that the evolved agents make use of the simpler embodied strategies first, as is shown from the evolved recurrent time-based network.

In summary, *a reactive controller in an embodied system doesn't imply reactive behaviour*: there is a difference between the local, instantaneous state definition of reactivity, and the behavioural definition, i.e., not being able to solve type-2 problems such as approach or avoidance under normal and inverted vision. As a result, whether an embodied agent will behave reactively (i.e., whether it will only be capable of performing behaviours of type-1) cannot be fully determined by the presence of a reactive controller.

The strategy proposed by minimally cognitive tasks for a critical examination of internal representation is straightforward: evolve agents on tasks that are 'representationally-interesting', then examine whether the agent is using representations [1]. In this case, no internal state is available for manipulation, thus, trivially, nothing is 'internally represented', yet behaviours such as commitment to discrete decisions on a categorisation task can still be demonstrated.

Future work will explore extensions to the capabilities of reactive controllers in a variety of directions. In general, it will be interesting to continue to relax *a priori* assumptions and consider how dynamical, bodily, and environmental constraints can transform 'cognitively hard' problems into easier ones. Some of these directions include: the effects of physical inertia, non-symmetrical architectures and noisy inter-node interactions.

References

1. Beer, R.D.: Toward the evolution of dynamical neural networks for minimally cognitive behavior. In Proc. of Simulation of Adaptive Behavior. MIT Press (1996).
2. Beer, R.D.: The dynamics of active categorical perception in an evolved model agent. Adaptive Behavior **11(4)** (2003) 209–243.
3. Chiel, H.J., Beer, R.D.: The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences* **20** (1997) 553–557.
4. Clark, A., Thornton, C. Trading spaces: computation, representation and the limits of uninformed learning. Behavioral and Brain Sciences **20** (1997) 57–90.
5. Dewey, J.: The reflex arc concept in physchology. Psychological Review **3** (1896) 357–370.
6. Di Paolo, E. A., Harvey, I.: Decisions and noise: the scope of evolutionary synthesis and dynamical analysis. Adaptive Behavior, **11(4)** (2003) 284–288.
7. Harvey, I.: The microbial genetic algorithm. Unpublished manuscript (1996).
8. Lashley, K., Schiller, C.: Instinctive behavior: the development of a modern concept. International University Press (1957).
9. Merleau-Ponty, M.: The Structure of Behavior. Beacon Press (1967).
10. Nolfi, S.: Power and the Limits of Reactive Agents. Neurocomputing **49** (2002) 119–145.
11. Pfeifer, R., Scheier, C.: Understanding Intellingence. Cambridge, MA. MIT Press (1999).
12. Scheier, C., Pfeifer, R., Kunyioshi, Y.: Embedded neural networks: exploiting constraints. Neural Networks **11** (1998) 1551–1569.
13. Varela, F., Thompson, E., Rosch, E.: The Embodied Mind: Cognitive Science and Human Experience. Cambridge, MA. MIT Press (1991).